PHD

Species-specific PCR primers for the rapid and reliable identification of yeast species

Harrison, Elizabeth

Award date:
2006

Awarding institution:
University of Bath

Species-specific PCR primers for the rapid and reliable identification of yeast species

Volume 1 of 1

Elizabeth Harrison

A thesis submitted for the degree of Doctor of Philosophy

University of Bath

Department of Biology and Biochemistry

June 2006

1st June 2009 E. Harrison

Contents

List of figures		i
List of tables		v
Acknowledgements		vi
Abstract		vii
List of Abbreviations		viii

				page
1	**Introduction**			**1**
	1.1	Definition of a species		1
	1.2	Traditional methods of yeast identification		2
	1.3	Molecular methods of yeast identification		3
		1.3.1	DNA base composition analysis	3
		1.3.2	DNA homology	4
		1.3.3	Ribosomal genes for identification	6
	1.4	Rapid yeast identification		10
		1.4.1	RAPD	11
		1.4.2	AFLP	12
		1.4.3	RFLP	12
		1.4.4	DGGE and TGGE	13
		1.4.5	Karyotype analysis	13
		1.4.6	Other methods	13
		1.4.7	Specific primers	14
		1.4.8	Advanced methods	15
	1.5	Bacterial species identification		15
	1.6	Fungal species identification		16
	1.7	Food spoilage yeasts		16
	1.8	Aim		20
	1.9	Strategy		21
		1.9.1	rDNA	21
		1.9.2	Housekeeping genes	22
		1.9.3	Species-specific genes	22
		1.9.4	Primer design strategy and parameters	22
2	**Materials and Methods**			**24**
	2.1	Yeast strains and culture conditions		24
	2.2	DNA preparation		24
		2.2.1	Cell number quantification	24
		2.2.2	Clean DNA preparation for PCR	24
		2.2.3	Rapid cell lysis for PCR	28
	2.3	Amplification		28
		2.3.1	Amplification with new primers	28
		2.3.2	Analysis of amplified PCR products	29
	2.4	Optimisation		29
		2.4.1	Temperature gradient PCR	29
		2.4.2	Magnesium ion gradient PCR	29
	2.5	Sequences		29
		2.5.1	Sequences from databases	29
		2.5.2	Sequencing primer design	30

2.5.3 Amplification for sequencing 32
2.5.4 Gel quantification of PCR products 32
2.5.5 PCR cleanup 33
2.5.6 Sequencing 33
2.5.7 Analysis of sequencing results 33
2.6 D1 / D2 identification of yeast species 33
2.7 Phylogeny 33

3 *Brettanomyces / Dekkera* **34**
3.1 *Brettanomyces* 34
3.1.1 Currently accepted species 34
3.1.2 Description of the genus 34
3.2 *Dekkera* 34
3.2.1 Currently accepted species 34
3.2.2 Description of the genus 35
3.3 History of *Dekkera / Brettanomyces* 35
3.4 Results 37
3.4.1 D1 / D2 sequence alignment 37
3.4.2 Specific primer design 37
3.4.3 Specific primer testing 40
3.5 Discussion 54
3.5.1 D1 / D2 sequence alignment 54
3.5.2 Specific primer design 55
3.5.3 Specific primer testing 55
3.6 Conclusion 60

4 *Zygosaccharomyces* **61**
4.1 Introduction 61
4.1.1 Currently accepted species 61
4.1.2 Description of the genus 61
4.1.3 Spoilage characteristics 61
4.1.4 History 62
4.2 Results 64
4.2.1 rDNA sequences 64
4.2.2 Housekeeping gene sequences 64
4.2.3 Specific primer design 76
4.2.4 Specific primer testing 78
4.2.5 Specific primer optimisation 84
4.3 Discussion 91
4.3.1 rDNA sequences 91
4.3.2 Housekeeping gene sequences 93
4.3.3 Specific primer design 96
4.3.4 Specific primer testing 96
4.4 Conclusion 103

5 *Saccharomyces* **105**
5.1 Introduction 105
5.1.1 Currently accepted species 105
5.1.2 Description of the genus 105
5.1.3 Spoilage characteristics 105

5.1.4 History 106
5.2 Results 108
5.2.1 Target sequences 108
5.2.2 Specific primer design 112
5.2.3 Specific primer testing 115
5.2.4 Multiplexing 147
5.3 Discussion 149
5.3.1 rDNA sequences 149
5.3.2 Specific primer design 149
5.3.3 Specific primer testing 150
5.3.4 *S. cariocanus* and *S. paradoxus* 150
5.3.5 *S. pastorianus* 151
5.3.6 Multiplexing 157
5.4 Conclusion 158

6 Concluding discussion 159
6.1 Primer characteristics 159
6.1.1 Total number of mismatches 159
6.1.2 3' mismatches 159
6.1.3 Primer melting temperatures 162
6.2 Genus specific primers 165
6.3 Conclusion 165

7 References 168

List of figures

Figure		page
1.1	The organisation of the yeast rDNA cistron	**6**
3.1	ClustalW alignment of the rDNA region D1 / D2 of the type strains of genus *Dekkera / Brettanomyces*	**38**
3.2	Maximum parsimony phylogenetic tree of the D1 / D2 sequences from the *Brettanomyces / Dekkera* type strain of each species	**37**
3.3	Target and closest sibling species amplified with all combinations of forward and reverse *Brettanomyces nanus* primers	**41**
3.4	*B. custersianus* strains amplified with combinations of the *B. nanus* primers	**42**
3.5	*D. bruxellensis* strains amplified with combinations of *B. nanus* primers	**42**
3.6	*D. anomala* strains amplified with combinations of the *B. nanus* primers	**42**
3.7	Target species amplified with all combinations of forward and reverse *B. naardenensis* primers	**43**
3.8	*B. custersianus* strains amplified with all combinations of forward and reverse *B. naardenensis* primers	**43**
3.9	*D. bruxellensis* strains amplified with all combinations of forward and reverse *B. naardenensis* primers	**44**
3.10	*D. anomala* and *B. nanus* strains amplified with all combinations of forward and reverse *B. naardenensis* primers	**44**
3.11	Species amplified with all forward and reverse combinations of *B. custersianus* primers	**45**
3.12	*B. custersianus* strains 1 and 2 amplified with all potentially *B. custersianus* specific combinations of forward and reverse primer	**45**
3.13	*D. anomala* strains amplified with all potentially *B. custersianus* specific combinations of forward and reverse primer	**46**
3.14	*D. bruxellensis* strains amplified with all potentially *B. custersianus* specific combinations of forward and reverse primer	**46**
3.15	*B. naardenensis* strains amplified with all potentially *B. custersianus* specific combinations of forward and reverse primer	**47**
3.16	*B. nanus* strains amplified with all potentially *B. custersianus* specific combinations of forward and reverse primer	**47**
3.17	Target species amplified with all combinations of forward and reverse *D. bruxellensis* primers	**48**
3.18	*D. anomala* strains amplified with all combinations of *D. bruxellensis* forward and reverse primer	**48**
3.19	*B. nanus* strains amplified with all combinations of *D. bruxellensis* forward and reverse primer	**49**
3.20	*B. custersianus* strains amplified with all combinations of *D. bruxellensis* forward and reverse primer	**49**
3.21	*B. naardenensis* strains amplified with all combinations of *D. bruxellensis* forward and reverse primer	**49**
3.22	*D. bruxellensis* and *B. naardenensis* strains amplified with all combinations of *D. bruxellensis* forward and reverse primer at 60°C annealing temperature	**49**
3.23	Target species amplified with all combinations of forward and reverse *D. anomala* primers	**50**
3.24	*D. bruxellensis* strains amplified with all combinations of *D.anomala* forward and reverse primer	**50**
3.25	*B. custersianus* strains amplified with all combinations of *D.anomala* forward and reverse primer	**50**
3.26	*B. nanus* strains amplified with all combinations of *D.anomala* forward and reverse primer	**51**

3.27 *B.naardenensis* strains amplified with all combinations of *D.anomala* forward and reverse primer **51**
3.28 *D. anomala* and *B. custersianus* strains amplified with *D. anomala* primers at different annealing temperatures **52**
3.29 ClustalW alignment of the rDNA region D1 / D2 of the *B. custersianus* type strain and strains 1 and 2 from this study **53**
3.30 Maximum parsimony phylogenetic tree of the D1 / D2 sequences from the *Brettanomyces / Dekkera* type strain of each species and *B. custersianus* strains **57**
3.31 *D. anomala* specific primer target loci in the closest sibling (*D. bruxellensis*) and the *B. custersianus* type strain and strains 1 and 2 **58**
3.32 A representative portion of the *B. custersianus* strain 1 D1 / D2 sequencing chromatogram readout **58**
4.1 ClustalW alignment of the *Zygosaccharomyces* species type strains rDNA region D1 / D2 **65**
4.2 ClustalW alignment of the *Zygosaccharomyces* species type strains rDNA region ITS1 **66**
4.3 ClustalW alignment of the *Zygosaccharomyces* species type strains rDNA region ITS2 **67**
4.4 ClustalW alignment of the *Zygosaccharomyces bailii* strains rDNA region ITS1 **68**
4.5 ClustalW alignment of the *Zygosaccharomyces bailii* strains rDNA region ITS2 **69**
4.6 ClustalW alignment of *Z. bailii* and *Z. rouxii LEU2* ORF **70**
4.7 ClustalW alignment of *Z. bailii* and *Z. rouxii HIS3* ORF **71**
4.8 Diagram showing the methodology used to sequence the 5' and 3' ends of the *Zygosaccharomyces HIS3* gene **73**
4.9 ClustalW alignment of the *Zygosaccharomyces species* partial *HIS3* gene **74**
4.10 ClustalW alignment of the *Zygosaccharomyces bailii* strains partial *HIS3* sequence **75**
4.11 Potentially *Z. bisporus* specific primers **78**
4.12 Examples of typical results when provisionally specific primer pairs are tested on all target isolates **79**
4.13 Examples of typical results when provisionally specific primer pairs are tested on their targets closest sibling **80**
4.14 Example of a provisionally specific primer pair amplifying unexpected species **80**
4.15 (A) *Z. bailii* specific primers F3 / R2 (B) *Z. bailii* non-specific primers F1 / R2 **82**
4.16 ClustalW alignment of the *Zygosaccharomyces species* and *L. cidri* partial *HIS3* gene **83**
4.17 Annealing temperature gradient PCR with *Z. lentus* primers F2 / R1 **85**
4.18 *Z. lentus* primers F2 / R1 at an annealing temperature of 65°C with target isolates and non-target isolates that are positive at 55°C **86**
4.19 Annealing temperature gradient PCR with *Z. bailii* primers F1 / R2 **87**
4.20 *Z. bailii* primers F1 / R2 with target isolates and non-target isolates that were positive at 55°C, at different Mg^{2+} ion concentration **88**
4.21 *Z. bailii* primers F1 / R2 with target isolates and non-target isolates that were positive at 55°C, at different Mg^{2+} ion concentration **89**
4.22 *Z. bailii* primers F1 / R2 with *Z. bailii* isolates under PCR conditions of 60°C annealing temperature and 1.8mM Mg^{2+} ions **89**
4.23 Annealing temperature gradient PCR with *Z. kombuchaensis* primers F3 / R1 **90**
4.24 A maximum parsimony phylogenetic tree constructed using the ITS1 sequences of the *Zygosaccharomyces* type species and *Z. bailii* strains **92**
4.25 A maximum parsimony phylogenetic tree constructed using the ITS2 sequences of the *Zygosaccharomyces* type species and *Z. bailii* strains **92**
4.26 A maximum parsimony phylogenetic tree constructed using the *HIS3* sequences of t he *Zygosaccharomyces* type species and *Z. bailii* strains **94**
4.27 A maximum parsimony phylogenetic tree constructed using the ITS1, ITS2 and *HIS3* sequences of the *Zygosaccharomyces* type species and *Z. bailii* strains **95**

4.28 The *Z. bailii* specific primer target loci in a variety of *Z. bailii* strains **97**
4.29 The *Z. kombuchaensis* specific primer target loci in the *Zygosaccharomyces* species **98**
4.30 The *Z. lentus* specific primer target loci in the *Zygosaccharomyces* species **99**
4.31 A maximum parsimony phylogenetic tree built using the D1 / D2 sequences of the type strains of clades 7, 8, 9 and 10 of the *Saccharomycetaceae* **100**
4.32 *Z. bailii* primer target regions in *Zygosaccharomyces sp.* and *L. cidri* **101**
5.1 ClustalW alignment of the *Saccharomyces sensu stricto* type species D1 / D2 rDNA region **109**
5.2 ClustalW amino acid alignment of the *Saccharomyces sensu stricto* species *FAL1* gene **114**
5.3 Maximum parsimony tree of *FAL1* gene sequences **115**
5.4 Potentially *S. mikatae* specific primer target loci in each of the *Saccharomyces sensu stricto* species **116**
5.5 Example of a typical specific primer pair result **117**
5.6 Example of a typical non-specific primer pair result **117**
5.7 ClustalW amino acid alignment of the *Saccharomyces sensu stricto* species partial *MEX67* gene **118**
5.8 Maximum parsimony tree of *MEX67* genes **119**
5.9 *S. cerevisiae* potentially specific primer target loci in each of the *Saccharomyces sensu stricto* species **120**
5.10 Potentially specific *S.cerevisiae* primer pair testing **121**
5.11 ClustalW amino acid alignment of the *Saccharomyces sensu stricto* species partial *PRI1* gene **122**
5.12 Maximum parsimony tree of *PRI1* genes **123**
5.13 *S. kudriavzevii* potentially specific primer target loci in each of the *Saccharomyces sensu stricto* species **124**
5.14 ClustalW amino acid alignment of the *Saccharomyces sensu stricto* species partial *VAS1* gene **125**
5.15 Maximum parsimony tree of *VAS1* genes **126**
5.16 *S. cariocanus* potentially specific primer target loci in each of the *Saccharomyces sensu stricto* species **128**
5.17 ClustalW amino acid alignment of the *Saccharomyces sensu stricto* species partial *SEC31* gene **129**
5.18 Maximum parsimony tree of *SEC31* genes **130**
5.19 *S. paradoxus* potentially specific primer target loci in each of the *Saccharomyces sensu stricto* species **131**
5.20 Partial D1/D2 sequence alignment of various putative *S.paradoxus* strains illustrating the single SNP between *S.paradoxus* and *S.cariocanus* **133**
5.21 Two portions of a ClustalW alignment of the D1 / D2 regions of the original "*S. paradoxus*" strains and the type strains of *S. paradoxus*, *S. cariocanus* and *S. cerevisiae* **134**
5.22 *S.cariocanus* strains 1 – 11 amplified with the *S.paradoxus* F1 / R2 primer pair at different annealing temperatures **135**
5.23 ClustalW amino acid alignment of *the Saccharomyces sensu stricto* species partial *DBP6* gene **136**
5.24 Maximum parsimony tree of *DBP6* genes **137**
5.25 *S. bayanus* potentially specific primer target sites in each of the *Saccharomyces sensu stricto* species **138**
5.26 ClustalW amino acid alignment of *the Saccharomyces sensu stricto* species partial *SPB1* gene **140**
5.27 Maximum parsimony tree of *SPB1* genes **141**
5.28 Representative results from testing the *S. pastorianus* potentially specific primer pairs on target isolates from the local collection **141**

5.29 Maximum Parsimony phylogenetic tree of the *Saccharomyces* type species, *S. bayanus* strains and *S. pastorianus* strains D1 / D2 rRNA gene **143**
5.30 ClustalW alignment of the D1 / D2 rDNA sequences of representative strains for each cluster shown in figure 5.29 **144**
5.31 Multiplex PCR with (A) All *Saccharomyces* species-specific primers in a single reaction mixture tested against the genomic DNA of each *Saccharomyces* species (B) *S. bayanus*, *S. mikatae* and *S. kudriavzevii* species-specific primers in a single reaction mixture tested against the genomic DNA of each *Saccharomyces* species **147**
5.32 Multiplex PCR with all *Saccharomyces* species-specific primers in a single reaction mixture, tested against different combinations of *Saccharomyces* species genomic DNA **148**
5.33 Maximum Parsimony phylogenetic tree of the *Saccharomyces* type species (barring *S. bayanus*), *S. bayanus* strains and *S. pastorianus* strains *FAL1* gene **153**
5.34 Maximum Parsimony phylogenetic tree of the *Saccharomyces* type species (barring *S. bayanus*), *S. bayanus* strains and *S. pastorianus* strains *MEX67* gene **154**
5.35 Maximum Parsimony phylogenetic tree of the *Saccharomyces* type species (barring *S. bayanus*), *S. bayanus* strains and *S. pastorianus* strains *PRI1* gene **155**
6.1 The total number of mismatches in the target loci of a primer pair compared to the proportion of positive and negative results with non-target species **160**
6.2 The melting temperature (°C) of the forward and reverse *Zygosaccharomyces* primers separated for positive and negative reactions **161**
6.3 The percentage of *Z. bailii* F1 primer bound to different targets at the experimental annealing temperatures displayed in figure 4.19 **163**
6.4 The percentage of *Z. bailii* R2 primer bound to different targets at the experimental annealing temperatures displayed in figure 4.19 **164**
6.5 ClustalW alignment of the *Zygosaccharomyces species* and *L. cidri* partial *HIS3* gene **167**

List of tables

Table		page
1.1	Group 1 spoilage yeasts	**19**
2.1	*Saccharomyces* strains used	**25**
2.2	*Brettanomyces / Dekkera* strains used	**26**
2.3	*Zygosaccharomyces* strains used	**27**
2.4	Other strains used	**28**
2.5	NCBI accession numbers of all downloaded sequences	**30**
2.6	Primers used to sequence *Saccharomyces* genes	**31**
2.7	Primers used to sequence rDNA	**32**
3.1	Sequences of potentially *Brettanomyces / Dekkera* species-specific primers	**39**
3.2	Expected amplicon sizes for all *Brettanomyces / Dekkera* potentially specific primer pairs	**40**
3.3	ClustalW distance matrix for the D1/D2 alignment in figure 3.1	**54**
4.1	Degenerate and specific primers that amplified the mid-section, 5' end and then 3' end of the *Zygosaccharomyces HIS3* gene	**72**
4.2	Primer sequences for amplifying and sequencing the *Z. bailii* complete *HIS3* gene	**73**
4.3	Expected amplicon sizes for all *Zygosaccharomyces HIS3* sequencing primers	**73**
4.4	Sequences of all potentially species-specific *Zygosaccharomyces* primers	**76**
4.5	Expected amplicon sizes for all *Zygosaccharomyces* potentially specific primer pairs	**77**
4.6	List of species in clades 8, 9 and 10	**81**
5.1	*Saccharomyces* lethal knockout genes	**110**
5.2	*Saccharomyces* target genes, systematic names, functional names and brief description of their function	**111**
5.3	Sequences of potentially *Saccharomyces* species-specific primers	**112**
5.4	Expected amplicon sizes for all *Saccharomyces* potentially specific primer pairs	**113**
5.5	The results with *S. kudriavzevii* primer pairs	**124**
5.6	The results with *S. cariocanus* primer pairs	**127**
5.7	List of original species assignation of "*S. paradoxus*" isolates that were tested with *S. paradoxus* potentially specific primers	**132**
5.8	The results with *S. paradoxus* primer pairs	**135**
5.9	The results with *S. bayanus* primer pairs	**138**
5.10	*S. pastorianus* isolates included in this study and their origins and phylogenetic history	**145**
5.11	Sequence variation in *S. pastorianus* strains compared to *S. bayanus* strains	**146**
5.12	A comparison of the PCR results using different primer pairs on *S. bayanus*, *S. cerevisiae* and *S. pastorianus* strains	**151**
6.1	The number of positive and negative results with non-target species when the 3' position is mismatched in both primers of a pair, one primer and no primers	**162**
6.2	The PCR results using the *Z. bailii* F1 / R2 primer pair and the total number of mismatches in the specific and non-specific target loci	**165**

Acknowledgements

I would like to thank Dr. Alan Wheals and Dr. Malcolm Stratford for their supervision and assistance and Dr. Hazel Steels and Dr. Steve James for their useful help and advice and the provision of strains. I would also like to thank the undergraduate project students that contributed to my work. Thanks also to members of lab 1.52 and room 1.55 past and present, particularly Kim Reilly and Carolyn Williamson for assistance and Hilton and Rob for the much appreciated light relief. I would finally like to thank my family and friends for help and support when I needed it.

Abstract

A novel method to identify yeast species using PCR primers is described. The aim was to produce a set of species-specific PCR primers that would rapidly and reliably identify target spoilage yeasts and their siblings using a simple, low-cost method that is ultimately suitable for high-throughput automation. Industrially significant spoilage yeasts of the *Brettanomyces*, *Dekkera*, *Zygosaccharomyces* and *Saccharomyces* species have been differentiated using this method. The method is suitable for well-defined, monophyletic taxa. The type of genomic sequence exploited by the PCR primers is dependent upon the taxonomic resolution of the target genus. The most important characteristic of each species specific primer pair is to encompass as many nucleotide polymorphisms between the target species and other yeasts. Generic PCR conditions have been used to amplify all strains of a target species despite intra-specific variation. Heterogeneous taxa such as *Zygosaccharomyces bailii* are successfully identified. In genera where small amounts of target sequence data are available, a single gene target is used for all species. Optimisation of PCR conditions is then required for specificity. In genera where large amounts of target sequence data are available, a different target sequence can be used for each species, allowing for multiplex reactions. *Saccharomyces pastorianus* has been investigated to assess the suitability of the method to identify hybrid yeasts and has shown that polyphyletic taxa cannot be identified using a single species-specific primer pair. The method has potential to also differentiate at the genus and sub-species levels.

List of abbreviations

AFLP	Amplified Fragment Length Polymorphism
B	*Brettanomyces*
BLAST	Basic Local Alignment Search Tool
CBS	Centraalbureau voor Schimmelcultures
CHEF	Contour-clamped homogeneous electric field
D	*Dekkera*
DGGE	Denaturing gradient gel electrophoresis
DMDC	Dimethyldicarbonate
ITS	Internal transcribed spacer
NTS	Non-transcribed spacer
L	*Lachancea*
LSU	Large subunit
mRNA	Messenger RNA
mtDNA	Mitochondrial RNA
NCBI	National Centre for Biotechnology Information
NCYC	National Collection of Yeast Cultures
nDNA	Nuclear DNA
NRRL	Northern Regional Research Laboratory
P	*Pichia*
PCR	Polymerase chain reaction
RAPD	Random amplified polymorphic DNA
PFGE	Pulsed-field gel electrophoresis
PNA	Peptide nucleic acid
rDNA	Ribosomal DNA
RFLP	Restriction fragment length polymorphism
rRNA	Ribosomal RNA
RT-PCR	Reverse-transcription PCR
S	*Saccharomyces*
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SGD	*Saccharomyces* Genome Database
SNP	Single nucleotide polyorphism
T	*Torulaspora*
TGGE	Temperature gradient gel electrophoresis
T_m	Melting temperature
tRNA	Transfer RNA
VNC	Viable but non-culturable
Z	*Zygosaccharomyces*
Zt	*Zygotorulaspora*

Chapter 1. Introduction

Yeasts are single-celled eukaryotes. Taxonomically, they are placed in the kingdom Fungi. The phyla Basidiomycota and Ascomycota form a discrete subgroup of the Fungi. Basidiomycota includes mushrooms, rust and smut fungi. Ascomycota includes sac fungi, yeasts and Penicillia and accounts for almost 75% of all described fungi. Ascomycota is separated into three subphyla: Archaeascomycetes, Euascomycetes and Hemiascomycetes. Archaeascomycetes include the fission yeasts, e.g. *Schizosaccharomyces.* Euascomycetes include morel, truffle and *Neurospora.* Hemiascomycetes include the budding yeasts, e.g. *Saccharomyces.* More than 1,000 yeasts species have been described to date [1].

Yeasts can be defined as fungi whose vegetative growth predominantly results from budding or fission and which do not form their sexual states within or upon a fruiting body. Asexual species were previously classified separately as Deuteromycota. Modern comparisons of nuclear DNA have shown some to be related to sexual yeasts. This has improved their taxonomy and allowed re-classification into Family "mitosporic Saccharomycetales" within the same Order as their sexual counterparts. For example, the sexual (teleomorphic or perfect) genus *Dekkera* has an asexual (anamorphic or imperfect) counterpart, *Brettanomyces.*

1.1 Definition of a species

A species is a group of closely allied mutually fertile individuals showing constant differences from allied groups. In organisms capable of sexual reproduction, species delimitation is established by applying the biological species concept; isolates are able to hybridise and produce viable and fertile offspring [2]. Yeast species identification has been relevant for many years and not only for taxonomic reasons. Identification is also useful for industrial purposes to identify the best brewing strains for beers and lagers and the best fermentation strains for various wines, vinegars and teas. In a medical context, the risk of mortality from yeast infection grows more prevalent in an increasingly immunocompromised population. The precise identification of the pathogenic species, and even strain, is essential to prescribe the most effective therapy. The antifungal drug, fluconazole is less effective against some *Candida* species and strains [3].

1.2 Traditional methods of yeast identification

Species definition based on interfertility is currently the only reliable criterion for exact species distinction. Even though it is technically ideal, it cannot be successfully applied to all strains. In some cases, strains isolated from natural and industrial environments do not sporulate and show anomalies in their ploidy. Difficulties encountered in applying the biological species concept mean that other differentiation methods are required. Traditional methods of identification began by distinguishing between species based on mating analyses to first define the phenotypic characteristics of a species. Following this, phenotypic tests were devised in an attempt to correctly assign species to new isolates. Nutrient assimilation profiles and morphological features were originally the only available quantifiable characteristics. Carbon and nitrogen assimilation patterns, as well as ascospore morphology and cell wall and envelope components were measured.

Nutrient assimilation profiles are established by growing a species with a variety of carbon and nitrogen compounds. Growth and absence of growth with each compound is recorded and compared against a database of known species nutrient assimilation profiles [4]. Other tests are also used to establish the broader taxonomy before exact species can be assigned. Ascomycetes and basidiomycetes differ in their modes of spore production. Ascomycetes form sexually produced spores within an ascus and basidiomycetes produce them on a basidium. The two types of fungi respond differently to diazonium blue B. Ascomycetes have no visible response while basidiomycetes give a reddish colour. Pyrophosphatidic acid is found exclusively in basidiomycetes [5]. Light microscopy and scanning electron microscopy (SEM) can also give a lot of information about a species and its taxonomy [6].

Analysis of the cell wall and envelope components for chemotaxonomy has been mainly directed at the mannan side chains of the ascomycetous cell envelope [7]. Significant differences in the carbohydrate moiety have been demonstrated between yeast species. Controlled acetolysis selectively cleaves the α-(1$\rightarrow$6) linkages of the mannan backbone. These are then separated according to size by gel filtration giving a reproducible characteristic elution pattern for a particular species, essentially a mannan fingerprint. Another approach to mannan fingerprinting was developed by Gorin and Spencer (1968) [8] who determined and compared proton magnetic resonance spectra of isolated mannans from a large number of yeast species. Spectra provided sensitive criteria for distinguishing mannans of different species or for showing their similarity.

1.3 Molecular methods of yeast identification

Morphological and physiological characters represent only a small portion of an organisms genetic potential. Analysing a species whole genome, or a representative portion of the genome, and comparing it with others allows a more accurate appraisal to be made.

1.3.1 DNA base composition analysis

DNA base composition analysis was the first method to directly compare genome characteristics in order to infer taxonomic relationships. This method gave the moles percent guanine plus cytosine value (GC content) of a genome by measuring either, a) the buoyant density of DNA extracted from an organism or b) determining the thermal denaturation temperature of a genome using a spectrophotometer. DNA base composition analysis was first used in bacterial taxonomy (reviewed in Marmur *et al* (1963) [9]). Storck (1966) [10] performed the first extensive study in fungi. Thirty species of fungi were examined including 14 species of yeast from 11 different genera. GC content of the ascosporogenous yeasts examined were located in a range of 39 to 45% GC. Stenderup and Bak (1968) [11] examined the DNA base composition of 18 species of *Candida*. The heterogeneity of the genus was exhibited by a wide range of GC values from 34.9 to 57.6%. Meyer and Phaff (1969) [12] examined the GC content of *Saccharomyces*, *Lodderomyces*, *Metschnikowia* and *Candida* species to assess their taxonomic position. From this they established closer relationships between some of the species examined.

Comparison of whole genomes via DNA base composition analysis gave only a superficial observation of the genetic potential of a species. Closely related organisms are expected to have similar base compositions although the inverse is not true [13]. This method only established if isolates linked via phenotypic characteristics were in fact the same species. To assess genome characteristics in full, a new method was required.

1.3.2 DNA homology

In 1970, Bicknell and Douglas [14] used DNA homology methods to estimate evolutionary divergence among species of the genus *Saccharomyces*. Genomic DNA of *S. cerevisiae*, *S. lactis* and *S. fragilis* (currently *Kluyveromyces lactis* and *K. marxianus* respectively) was filter bound to a membrane and annealed with radiolabelled homologous DNA or LSU (large subunit) rRNA. This was then competed off with unlabelled heterologous DNA or LSU rRNA, and the percent relatedness between species was measured. It was found that there was little homology within the *Saccharomyces* group (this is not surprising as today we know that the *Saccharomyces* group of 1970 included a broad array of yeasts that are now considered to be in separate genera).

At this stage, molecular yeast taxonomy was limited to a small collection of investigations on a diverse variety of yeast groups. It took advantage of the wealth of physiological and morphological data that had been previously gathered, as well as results from mating experiments. Reinforcing this with genome comparison methods allowed speciation conclusions to be drawn. However, no generic molecular method to conclusively identify a species had been found. Price *et al* (1978) [15] reviewed the DNA homology methods that had been used. They determined the DNA base sequence similarity among phenotypically similar species of *Schwanniomyces*, *Saccharomyces*, *Debaryomyces* and *Pichia* to establish if existing taxonomic criteria reflected the true evolutionary affinities within the group. By correlating the molecular results with morphological and physiological results, Price *et al* attempted to infer generic principles to molecular taxonomy. They found that as a consequence of the limitations of the conventional system, many of the yeast species described in the literature represented an amalgam of distantly related organisms, while other species had been separated on trivial grounds from closely related taxa. It was suggested that 80 to 100% DNA relatedness constituted a species. On continuing to develop the process of molecular identification, they concluded, "Considerable information concerning the natural relationships among these yeasts should result from additional comparison of their informational macromolecules."

A lack of ways to efficiently compare the informational macromolecules of species, lead to a variety of techniques being used to classify groups of species into more accurately delimited genera. In 1980, Kurtzman *et al* [16] used a multifaceted approach involving ascospore morphology, mating reaction and DNA reassociation to reorganise *Pichia* species. They found that *P. kudriavzevii*; *P. terricola* and *P. scutulata* formed a distinct subgroup when ascospore ultrastructure was examined; they had warty ascospore walls while other *Pichia* were smooth. DNA reassociation values and mating reactions agreed that these warty *Pichia* species formed a distinct subgroup and they were reassigned to genus *Issatchenkia* Kudriavzev. It was concluded that in agreement with Price *et al* (1978) [15], a DNA relatedness value between 80 and 100% constituted conspecificity. Kurtzman *et al* (1980) [17] investigated other *Pichia* species using DNA reassociation and mating reactions. Two new *Pichia* isolates collected from Mississippi pine trees differed from other species in starch assimilation. Genetic crosses were undertaken and frequency of conjugation was seen to be comparable to that for intraspecific pairing. However, most of the ascospores resulting from the crosses were poorly formed. The few spores isolated were not viable. This would indicate that the new isolates were separate species but closely related. DNA reassociation results showed that the species exhibited only 25% DNA relatedness, suggesting they were not as closely related as first thought. Varieties of *Pichia scutulata* had similar DNA reassociation values but produced viable ascospores at a low frequency. Kurtzman *et al* concluded, "defining species strictly on the basis of a certain percentage of DNA relatedness may conflict with the biological reality of genetic exchange. Almost certainly, considerable latitude will be found in the definition of species even at the genetic and molecular levels." It became evident that DNA homology values alone were not sufficient to conclusively delimit species.

The main limitation of DNA homology studies is that only information on pairs of species is given. Each species must be experimentally tested against everything else if definite conclusions are to be drawn. Although the methodology combined with classical techniques is good for assigning species, it can only infer broad associations between species and in yeast taxonomy as a whole. Additional information is required for complete resolution among species and genera.

1.3.3 Ribosomal genes for identification

Over the next few years, there were increasing numbers of investigations into the use of the ribosomal DNA in yeast taxonomy. Because the function of ribosomes are to translate genes into protein, they are ubiquitous in all living organisms. As such, their sequence is highly conserved so as to retain a function that is essential for life. The functional ribosome is comprised of two protein and RNA subunits, one large and one small. The small subunit binds the mRNA and tRNA, while the large subunit catalyses peptide bond formation. More than half of the ribosome is RNA. In eukaryotes, the 18S rRNA combines with proteins to make the small ribosomal subunit. 5S, 5.8S and large subunit (LSU) rRNA plus additional proteins make the large ribosomal subunit. The fungal LSU can be between 25 and 28S (where S equals Svedberg units, a measure of the rate of sedimentation of a particle in a centrifuge, where the sedimentation rate is associated with the size of the particle).

The eukaryotic rRNA genes are encoded in a tandemly repeated cistron composed of the 18S, 5.8S and 26S genes [18]. Internal transcribed spacer (ITS) regions separate the subunits. Non-transcribed spacer (NTS) regions separate the cistrons (figure 1.1). In *S. cerevisiae*, the 5S open reading frame is on the strand opposite the NTS [19], but it can also be present elsewhere in the genome of lower eukaryotes. Higher eukaryotes organise the 5S genes into separate clusters. Between organisms, the structure of the rRNA gene cluster is highly conserved but numerous differences occur with respect to copy number, structure of the NTS and homogeneity of the repeat unit.

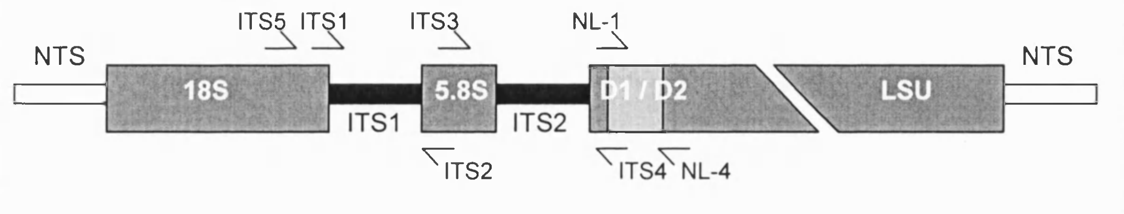

Figure 1.1: The organisation of the yeast rDNA cistron and the position of primers for amplifying various subunits. ITS primers are from White *et al* (1990) [20], NL primers from Kurtzman and Robnett (1997) [29].

The first extensive use of rRNA comparisons for yeast systematics was the DNA homology investigation by Bicknell and Douglas (1970) [14] described previously. White *et al* (1990) [20] published a set of primers that were conserved in all fungi and could amplify individual subunits, or the complete rDNA cistron, using PCR. The primers made this part of the genome of all yeasts and fungi accessible for routine amplification, sequencing and comparison. The ribosomal DNA became the subject of multiple investigations as to its suitability for improving efficiency and resolution in molecular yeast taxonomy. Lachance (1990) [21] restriction mapped the rDNA repeat unit of 125 isolates of the cactophilic yeast species, *Clavispora opuntiae* with 12 restriction enzymes. He explored the significance of rDNA variation of a worldwide collection to investigate if patterns were indicative of biogeographic, ecological or microevolutionary aspects of that yeast.

As molecular biological methods improved in efficiency and sophistication, more techniques were brought in to compare specific parts of the genome. In 1987, Magee *et al* [22] used restriction fragment length polymorphism (RFLP) analysis of the ribosomal repeat DNA of *Candida* species to assess whether the restriction map of the rDNA inferred phylogenetic relationships. Clark-Walker *et al* (1987) [23] used the mitochondrial DNA to assess relationships between *Brettanomyces*, *Dekkera* and *Eeniella* species. They concluded that some of the *Brettanomyces* species were anamorphs of *Dekkera* species.

Peterson and Kurtzman (1991) [24] assessed a variety of regions in the rDNA for suitable amounts of variability to consistently distinguish between yeast species. They looked at 4 regions of the 18S and 2 regions of the large subunit rDNA. The test was performed on well-defined isolates. All of the regions tested were identical between species apart from one, 25S-635 (named after the starting nucleotide position in *S. cerevisiae*). This region had nucleotide polymorphisms between all species investigated but was identical between isolates of the same species. It was potentially a target for accurately distinguishing between all yeast species. *Saccharomyces cerevisiae* and *Saccharomyces bayanus* were separated by 5.4% difference. The most divergent species tested: *Saccharomyces cerevisiae*, *Schizosaccharomyces japonicus* and the basidiomycete *Leucosporidium scottii*, were all separated by the same degree of difference, showing that there are structural and functional constraints on the rRNA molecule. This limits its use as an evolutionary chronometer for distant relationships.

In 1992, Kurtzman [25] reviewed the progress of yeast taxonomy resulting from rDNA sequence comparisons and the future potential. He pointed out that "[ribosomes provide] a genetic map that charts the positions of all living things". Some parts of the rDNA sequence are sufficiently conserved that they are homologous for all organisms and serve as reference points for the alignment of less conserved areas for measurement of changes over varying evolutionary distances. Different regions of the rDNA exhibit different rates of nucleotide substitution; they are essentially a collection of chronometers. Different taxonomic contexts require different degrees of "phylogenetic resolution" in the organismal trace [26]. Regions that have a low rate of substitution can be used to assess distant relationships. Regions with higher substitution rates can assess closer relationships. Kurtzman concluded, "While it is still early, the studies that have been completed suggest that comparisons of rRNAs will significantly influence our perception of yeast systematics and its implicit reflection of evolution among fungi."

Fell *et al* (1992) [27] assessed the LSU to establish a technique for routinely identifying marine eukaryotes. They found that not only could synonyms be identified, but degrees of relatedness could also be established by counting the number of differences in the sequence.

James *et al* (1996) [28] assessed the use of different rDNA regions for organising the *Zygosaccharomyces* and *Torulaspora* genera. They found that among these species, the 5.8S was highly conserved. The 18S and ITS sequences suggested that the two genera were phylogenetically intermixed. The ITS sequence can be used to assess more closely related species due to its increased heterogeneity. Its hypervariability prevents its use for assessing wider species relationships. James *et al* concluded that a variety of rDNA sequences should be used when trying to accurately assess phylogenetic relationships.

Kurtzman and Robnett (1997) [29] sequenced the variable region of the LSU (named D1 / D2 as the area encodes the D1 and D2 loops in the secondary RNA structure of the ribosome) of all recognised *Candida* species using the primers NL-1 and NL-4 (table 2.7). They compared the sequence data with nuclear DNA complementarity results to assess the usefulness of the technique. Species with as little as 40% nDNA complementarity had identical D1 / D2 sequences. It had previously been accepted that nDNA complementarity values of 80 – 100% constituted identical species. These values were obviously too high. Members of a species showed 0 – 2 nucleotide differences in the region. They concluded that, "[it] appears to be possible to accurately identify most yeast species based on their nucleotide sequence in LSU rDNA region D1/D2."

James *et al* (1997) [30] used 18S sequence comparisons of all *Saccharomyces* yeasts to assess the convoluted phylogeny of this genus. They found that the genus was polyphyletic. A group of *Saccharomyces* were tightly related and formed the *Saccharomyces "sensu stricto"* group. The rest of the genus was highly diverse and formed the *Saccharomyces "sensu lato"* group.

Belloch *et al* (1998) [31] used restriction mapping to assess the ITS and 5.8S of *Kluyveromyces* species. This was compared to the phylogenetic tree derived from mitochondrial *COXII* gene sequences (Belloch *et al* (2000)) [32]. The trees were mostly congruent and showed the genus to be polyphyletic with one monophyletic group of species that could be considered the "true" *Kluyveromyces*. Cappa and Cocconcelli (2001) [33] used 18S sequencing to identify species from spoiled dairy products.

Kurtzman *et al* (1998) [34] sequenced the D1 / D2 region of all ascomycetous yeasts with primers NL-1 and NL-4. They compared the sequencing results with nuclear DNA relatedness species delimitations and found that, "nearly all currently recognized ascomycetous yeasts can be identified from their unique D1 / D2 sequences. Consequently, use of this database has the potential to markedly increase the accuracy of yeast identifications." From this point onwards, D1 / D2 sequence analysis became the fastest and most simple method of reliable species identification.

With differing areas of the genome producing slightly different phylogenies due to varying nucleotide substitution rates, it was important to know how much information was required to produce completely resolved results. Rokas *et al* (2003) [35] assessed how much concatenated sequence data was required to recover fully resolved phylogenetic trees with maximum support on all branches. They found that the sequence of at least 20 genes were required. It could be considered that this amount of sequence data is the only way in which to derive completely accurate phylogenies, although less data can infer adequate ones. As sequencing becomes more rapid and cost effective and the number of complete genomes elucidated increases, this will become a more feasible way of conclusively deriving phylogenies. However, it is not a practical way to rapidly identify species.

As molecular and genome survey techniques have developed, species identification methodologies have evolved and so improved the resolution of taxonomy. At this point, we can now empirically define a species by observing the degree of difference in a pre-defined area of the genome in different species

1.4 Rapid yeast identification

In industrial, environmental and medical settings, it is essential that contaminating or infecting species be identified swiftly so that the correct course of remedial action can be undertaken as soon as possible. It is important for the method to be rapid, cost effective, efficient and completely specific, particularly for non-sterile samples where a pathogen, spoilage or indicator organism must be determined in the presence of other organisms. Despite the consistent accuracy of D1 / D2 sequencing for species identification, it is still a relatively slow and expensive method. Sequencing is a sensitive technique and sample preparation should be accurate. Reactions often have to be outsourced to specialist sequencing laboratories, increasing the time until the isolate is identified. As molecular techniques become more sophisticated, researchers have been trying to develop more rapid methods of species identification.

D1 / D2 sequence evaluation has helped to delimit species and genera and has enabled the construction of accurate phylogenies. With intra- and inter-species relationships now fully resolved for some groups, these phylogenies have provided a basis for establishing new rapid identification methods to fulfil industrial and medical needs to minimise economic losses and pathogenicity.

The advent of the polymerase chain reaction (PCR) [36] has caused an increase in the number of molecular identification techniques available. This is due to the inherent rapidity and simplicity of the procedure. PCR amplifies any region of the genome depending upon the target of the oligonucleotide primers. The genetic material generated can then be visualised on an agarose gel allowing the researcher to progress from minute amounts of material to visual results within a matter of hours. The most prevalent molecular methods for species identification are random amplified polymorphic DNA (RAPD) analysis, amplified fragment length polymorphism (AFLP) analysis, restriction fragment length polymorphism (RFLP) analysis and gel methods, such as denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE). Other methods include karyotype analysis using contour-clamped homogeneous electric field (CHEF); total protein analysis using sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and species-specific peptide nucleic acid (PNA) probes.

1.4.1 RAPD

Random amplified polymorphic DNA analysis utilises PCR primers targeted to arbitrary sequences or various repeating elements throughout the genome. Such elements include microsatellites and δ elements. δ elements are direct repeat elements that flank the TY1 yeast retrotransposon. Microsatellites form a distinct class of repetitive sequences, characterised by direct tandem repeats of a short DNA motif, usually less than 10bp [37]. The same oligonucleotide sequence is used for both the forward and reverse primer. Amplifiable distances separate some of the repeats and the result of a PCR is a fingerprint of products. Some primers have been shown to give species specific fingerprints with certain genera. Comparison of fingerprints allows species to be assigned. A combination of two microsatellite primers identified yeast species in dairy products [38]. Microsatellite RAPD was used to identify *Zygosaccharomyces* species in a food production line. Strain discrimination allowed the origin of spoilage to be traced [39]. Microsatellite fingerprinting was also used to type 77 clinical *S. cerevisiae* isolates and found considerable interspecific diversity [40]. A combination of RAPD and RFLP analysis identified *Saccharomyces* and *Zygosaccharomyces* species [41].

Strain specific fingerprints are also possible. This method has been used to type brewery and wine yeast strains [42, 43, 44, 45]. The long terminal repeat in a putative yeast TY1 retrotransposon has also been used to characterise *Pichia membranefaciens* strains [46]. A multiplex method that simultaneously targeted three microsatellites successfully discriminated between *C. albicans* clinical strains [47].

An inherent problem with RAPD analysis is reproducibility. It can be difficult to obtain the same results under identical conditions in different laboratories because it relies on amplification from primers where the matches are often imperfect. Weak bands may disappear with alternative reagents and thermocyclers. Fingerprints can also be quite complex and it may be difficult to distinguish between subtly different patterns.

1.4.2 AFLP

Amplified fragment length polymorphism analysis involves the selective PCR amplification of restriction fragments from a total digest of genomic DNA [48]. The analysis requires three steps, (i) Restriction of total DNA and ligation of oligonucleotide adapters, (ii) Selective amplification of sets of restriction fragments using primers targeting the adaptors and restriction sites, (iii) Denaturing polyacrylamide gel analysis. Selective amplification can be achieved by extending the primer targets into restriction fragments. Between 50 and 100 restriction fragments can be detected depending on the gel resolution. This is a novel and powerful fingerprinting technique for DNAs of any origin or complexity and has been successfully used to type commercial, type and winery yeasts, including *Saccharomyces* and *Brettanomyces* [49] and strains used to produce various Mexican alcoholic beverages [50]. The advantages of this method are reproducibility and a potential for application across all phyla. However, it is initially labour intensive and is unable to establish genetic similarities [49]. The fingerprints produced are also highly complicated.

1.4.3 RFLP

Restriction fragment length polymorphisms assign species on the basis of the pattern of restriction products given when a region of the genome is amplified and cut with a restriction endonuclease, or combination of endonucleases. Typical targets include mitochondrial DNA (mtDNA) and rDNA.

RFLP analysis has been used to monitor the species at various stages of wine fermentation [51, 52, 53, 54, 55, 56, 57], to identify the species causing spoilage in yoghurt [58], to characterise wild *Zygosaccharomyces* species [59], to identify opportunistic pathogenic *Malessezia* species [60], to characterise the yeasts related to surface ripened cheeses [61], to identify *Saccharomyces* and *Torulaspora* species [62] and to delimit brewery yeast strains [63] in combination with RAPD analysis [64].

RFLP is a difficult method to establish. A variety of enzymes need to be tested on known species before a suitable one or combination can be chosen for discrimination of species. These enzymes may then only be suitable for a small number of species. The fingerprints produced can be difficult to interpret and a single nucleotide polymorphism (SNP) within an enzyme target site can prevent it from being cut. The addition of a restriction step between amplification and electrophoresis adds time and cost to the method. SNPs occur between strains of the same species and so some enzymes can be used to produce strain specific profiles.

Fingerprinting methods are a useful form of rapid identification. However, they have a number of drawbacks. The fingerprints can be very difficult to interpret due to the number of bands produced on a gel and similarity in size. They can also have high levels of intra-species variability. This is useful for identifying strains but not if species-level identification is required. There is also no definitive fingerprinting method that is useful for all yeasts. Different labs use different techniques that cannot be easily and reproducibly exchanged.

1.4.4 DGGE and TGGE

Denaturing gradient gel electrophoresis (DGGE) separates species according to the GC content of a particular area of the genome. A portion of the genome is amplified by PCR and then a denaturing gradient in the gel denatures the DNA duplex depending on melting temperature (which depends on GC content) [65]. Most often, rDNA regions are amplified in a PCR and then assessed using DGGE to assign species. The D1 region of the 26S rDNA was used to profile the yeasts involved in wine fermentations [66] and the partial LSU rDNA has also been used to study the yeast population dynamics during cocoa fermentations [67] and sourdough fermentations [68]. Temperature gradient gel electrophoresis (TGGE) of the ITS2 sequence has been successfully used to distinguish between enological *Saccharomyces sensu stricto* species and strains [69]. TGGE of the D1 / D2 regions has been successfully used to monitor yeast populations in pickle soaking fluid [70].

These methods are useful for directly sampling mixed populations but are also technically demanding.

1.4.5 Karyotype analysis

Karyotype analysis can be performed by pulsed field gel electrophoresis (PFGE) [71] or contour-clamped homogeneous electrical field (CHEF) [72, 73]. Chromosomes are separated on the basis of size and then compared for species identification. This kind of analysis is particularly useful for the identification of hybrids [74]. However, both methods are time consuming at the chromosome preparation and electrophoresis stages.

1.4.6 Other methods

SDS-PAGE of proteins – Total protein electrophoretic profiles have been used to distinguish between brewing yeasts [72]. However, the large number of bands and the closeness of these bands require expert analysis of the gels. As it is an analysis of genetic products, a rigorous standardisation must be applied to reduce the influence of the environment on gene expression. Species cannot be isolated from mixed cultures. A selection of enzymes can be used instead of total proteins. Enzymes commonly compared include hexokinase and alcohol dehydrogenase amongst others [75].

PNA probes – Peptide nucleic acid (PNA) is a pseudo peptide that strongly and specifically binds to nucleic acids. PNA probes have been used to target specific rDNA sequences *in situ*, which are then detected using a chemiluminescent reaction [76]. This method has an advantage of enumerating the filter bound organisms although it is costly.

Fatty acid composition [77] – Fatty acid composition analysis has been shown to differentiate fruit juice spoiling yeasts according to their technological significance. Fatty acid profiles can be used as zymological indicators, rather than conclusive species indicators. Although this is a useful trait, the importance of rapidly identifying preservative resistant yeasts such as *Z. bailii* is a primary concern.

Most often, the above methods are used in various combinations and with classical techniques in order to achieve complete species and/or strain identification. A combination of D1 / D2 sequence analysis, 5.8S RFLP and classical and phenotypic methods were used to identify yeast species in orange juice [78]. CHEF karyotype analysis, SDS-PAGE total protein analysis and arbitrary primer RAPD were used to characterise Brazilian fuel alcohol strains [72]. Karyotype analysis combined with mtDNA RFLP and δ sequence RAPD were used to type commercial *Saccharomyces* strains [53]. ITS and mtDNA RFLP in conjunction with ITS size polymorphism were used to identify yeasts associated with surface ripened cheese [61]. Karyotype analyses, RFLP of the *MET2* gene and growth tests were used to identify *Saccharomyces sensu stricto* yeasts from wine fermentations [79].

1.4.7 Specific primers

In the 1992 review by Kurtzman [25], the progress in yeast taxonomy and identification resulting from rDNA sequence comparisons was assessed. He concluded, "It is already apparent that a practical application of the work will be the development of taxon-specific oligonucleotide probes for rapid identification of medically, agriculturally, and industrially significant organisms". James *et al* (1996) [28] noted that there were conserved regions in the ITS sequences that can be used for species delineations and that, "such regions could also be invaluable for rapid species identification with probes". From an early time, it was obvious that PCR and the sequence differences between species could be combined to devise species-specific probes for identification. The most accurate way to rapidly identify a species is to design PCR primers that will amplify only the species you are looking for and nothing else. Such primers have been previously targeted to the intron splice sites [80, 81], ITS regions of *Brettanomyces* and *Dekkera* species [82] and *Trichosporon asahii* [83], *cs1* gene of *Candida krusei* by reverse transcription (RT)-PCR [84], *ACT1* exon and intron sequences of *Candida dubliniensis* [85], the pH-regulated *PHR1* gene of *C. albicans* [86] and the 26S rDNA of various yeasts [87, 88]. Methods not involving PCR but still using specific oligonucleotides include fluorescent *in situ* hybridisation of 18S specific probes in yoghurt spoiling yeasts [89]. These have all targeted single species or a very small group of species and have not provided a universal method that can be applied to all yeast species.

PCR using nested primers is another method for identifying species. A pair of generic PCR primers first amplifies a conserved region within the genome, which then provides the target for specific primers. 18S nested PCR was used to distinguish between fungal species causing endophthalmitis [90]. The ITS1-5.8S-ITS2 region was successfully used to distinguish between pathogenic *Candida* species in serum samples [91].

1.4.8 Advanced methods

Technological advances have resulted in new, sophisticated methods being used for species identification. Such methods include the use of real-time RT-PCR of the actin gene sequence specific primers to identify viable fungi involved in pasteurised food spoilage [92]. TaqMan quantification PCR has been used to specifically quantify pathogenic *Candida* cells in water [93]. Melting peak analysis of an rDNA amplicon has been used to distinguish between the spoilage yeasts in fruit juices [94].

1.5 Bacterial species identification

Bacterial molecular identification is currently more advanced than in yeast because pathogenicity has resulted in more research and genome sequencing. Genomics can identify a gene that is present in one species and absent in siblings. Specific recognition of that sequence using one of a variety of methods results in identification of the species. *Escherichia coli* strains have been amplified using species-specific primers to the *E. coli* specific Shiga-like toxin genes *stx1* and *stx2* and melting curve analysis to determine if strains contain one gene or the other, or both [95]. Lyme disease causing spirochetes, *Borrelia* species have been identified and distinguished between using specific plasmid sequences [96].

More advanced methods of bacterial species-specific identification are now being researched. Due to bacterial pathogenicity, quantity in food is important for legislative purposes. *Staphylococcus aureus* has been quantified from a variety of cheeses in real time using melting curve analysis of the *nuc* gene in a LightCycler [97]. *Campylobacter* species have been specifically amplified using genus-specific 16S primers and biprobes to distinguish between species [98]. *Borrelia* species have been specifically identified using genus-specific p66 primers and melting curve analysis [99]. *C. jejuni* species have been specifically quantified from water after DNase treatment to be sure of only amplifying live or viable but non-culturable (VNC) cells [100].

1.6 Fungal species identification

Fungal molecular identification is only at a basic stage. Most methods are confined to the ribosomal DNA. Specific PCR primers have been successfully targeted to the ITS region of *Fonsecaea pedrosi* [101], *Spongospora subterranea* f. sp. *subterranea* [102], *Rhizoctonia solani* [103], *Fusarium solani* [104] and *Pythium* species [105]. A LightCycler method has been used to identify dermatological fungal species by amplification and melting peak analysis of the ITS regions [106].

1.7 Food spoilage yeasts

Yeasts are known to spoil fruit juices and their concentrates, soft drinks and carbonated soft drinks. The typical chemical composition of these products is low pH and fermentable sugar content. Fruit juices and soft drinks contain ca. 10% sugar, principally fructose, glucose and sucrose. Oxygen levels are generally low in these products and are further minimised by the use of hot filling, pasteurisation and antioxidants. Soft drinks and fruit juices are technically acidic broth media. Low pH strongly suppresses bacterial growth and so fermentative yeasts dominate the resulting microbial flora. Molds and acid tolerant bacteria grow only where oxygen levels are sufficient.

Preservation of these products can be divided into physical and chemical methods. Physical methods involve maintaining a low pH and low water activity, pressure and carbonation, pasteurisation and chilled storage. Chemical methods include avoiding the preferential sugars of spoilage organisms, addition of acidulants, limiting microbial nutrients and oxygen in a product and the action of natural inhibitors from fruit, and preservatives. Preservatives come in the forms of weak-acids, parabens and dimethyldicarbonate (DMDC).

Throughout the manufacturing process, products are vulnerable to infection. At the bottling stage, microbial stability is established via a few methods: in-pack sterilisation; in-line sterilisation and aseptic filling; in-line pasteurisation followed by hot filling or addition of preservatives; or solely the addition of preservatives. The main off-site source of infection arrives with the raw materials used to make the product. Fruit is the major natural habitat for many yeast and mold species. It is regarded as routinely heavily contaminated; fruit juices from healthy fruit may contain $10^3 - 10^5$ CFU ml^{-1}. Those from mummified or surface damaged fruit can be several orders of magnitude higher than that. The natural yeast flora of fruit therefore includes fructophilic and osmotolerant yeasts, amongst other species. These yeasts are ideal spoilers of drinks with high sugar content and the production plants are also ideal environments for the yeasts to establish themselves. Manufacturing machinery is a source of product contamination if it contains colonies of such species. Sugar is another raw material that can be a source of contamination. Dry, granulated sugar does not support many microbes, however soft drinks plants often receive their sugar in a concentrated syrup format that can support osmotolerant yeasts. Other sources of microbes include aerial contamination that can be dependent on geographic location. Breweries and wineries nearby can affect the composition of the aerial flora, particularly with increased levels of *Saccharomyces cerevisiae*. This yeast is ubiquitous and entirely domesticated with no known natural habitat. It is absent in soft drinks factories in Iraq, a country with no alcoholic beverage production.

Insects are the most important aerial vectors for yeasts. Fruit flies are attracted to soft drinks, fruit juices and fermented beverages. *Drosophila* consumes many types of yeast and exhibit a preference for certain species. Packaging and additives are other sources of contamination.

Spoilage is caused by microbial growth and metabolism, as opposed to the incidental presence of non-growing organisms. Metabolism can result in clouds, sediments or particulates in a product. Taints may be caused by metabolites produced by actively growing yeasts and molds. Most spoilage yeasts are highly fermentable forming carbon dioxide and ethanol from sugars. Ethanol imparts a sweet taste but the biggest danger is caused by increased pressure due to gas production. Enough pressure can be generated to bulge cartons, split cans and cause explosions in glass or plastic bottles and aluminium kegs. Microbial splashes on other packaging can render large consignments unsaleable.

Non-spoilage organisms present in a product are a measure of hygiene in production. These should not be viewed with complacency, as alterations in production can allow adventitious growth.

Spoilage organisms are separated into four groups according to their order of importance. Group 1 contains species that cause spoilage. Group 2 contains marginal spoilage species and hygiene indicators. Group 3 contains only hygiene indicators and group 4 contains aliens; microbes that are out of place in that environment. The members of group 1 are listed in table 1.1.

The risks to human health from yeast spoilage of soft drinks and fruit juices are relatively minor. However, there is significant risk of injury from exploding bottles, particularly eye injury. Yeasts are generally unlikely to have adverse effects on human health although some are regarded as human pathogens. They are almost exclusively the concern of immunocompromised patients. These yeasts are *Candida famata*, *C. guilliermondii*, *C. krusei*, *C. parapsilosis* and *S. cerevisiae* [107].

Group	Species
1	*Dekkera anomala*
1	*Dekkera bruxellensis*
1	*Dekkera naardenensis*
1	*Hanseniaspora uvarum*
1	*Kluyveromyces marxianus*
1	*Kluyveromyces thermotolerans*
1	*Lodderomyces elongisporus*
1	***Zygosaccharomyces bailii***
1	***Zygosaccharomyces bisporus***
1	***Saccharomyces cerevisiae***
1	***Saccharomyces exiguous***
1	***Zygotorulaspora florentinus***
1	***Saccharomyces kluyveri***
1	***Torulaspora microellipsoides***
1	***Lachancea fermentati***
1	***Torulaspora delbrueckii***

Table 1.1: Group 1 spoilage yeasts. Species in bold are noted spoilage yeasts

1.8 Aim

The aim of this project was to produce a set of species-specific PCR primers that would rapidly and reliably identify target spoilage yeasts and their siblings using a simple and low-cost method. All other molecular methods require a period of analysis after the PCR reaction if not subsequent other steps, such as fingerprint interpretation, restriction or high-resolution electrophoresis steps. This method will be a simple band / no band result depending on the presence or absence of a target species. An empty lane on the gel will indicate the species is not present. A single band of pre-defined size will indicate that it is present. When compared with RAPD or RFLP techniques, it will be simpler to interpret, enabling the procedure to be undertaken by minimally trained staff in a factory. The only equipment required will be a PCR machine, agarose gel electrophoresis apparatus and visualisation equipment. The method will also be suitable for automation and high throughput using a real-time PCR machine that does not need an electrophoresis step prior to analysis.

1.9 Strategy

The genera that will be targeted with species-specific primers in this project are *Brettanomyces / Dekkera*, *Saccharomyces* and *Zygosaccharomyces.*

The basis of species-specific PCR primers takes advantage of sequence differences between a target species and its closest siblings. It also works on the assumption that more divergent species will have increased differences in the primer-binding site and will not be amplified. The reorganisation of the *Zygosaccharomyces* genus into four genera: *Zygosaccharomyces*, *Zygotorulaspora*, *Torulaspora* and *Lachancea* [133], has not changed the fact that these species have similar physiological characteristics making them good spoilage organisms. For this reason, the species that were classed in the *Zygosaccharomyces / Torulaspora* Superfamily of spoilage yeasts are considered siblings for the purpose of this work. This is because their physiological characteristics mean that they may well be present in the same environment and so a species-specific primer for one *Zygosaccharomyces* species must be proven to exclude the other members of this spoilage Superfamily. While molecular taxonomy draws attention away from physiological characteristics and onto ribosomal sequence data, it cannot be ignored that the physiological attributes of an organism determine its environment.

The design of PCR primers requires sequence data to be known in the target species and preferably all of its closest siblings. Different parts of the genome may be exploited for this method depending on the relationships between a species and its siblings.

1.9.1 rDNA

The most obvious target sequence data is the rDNA. It is already used to distinguish between species, at least the D1 / D2 sequence of all known species has already been obtained and primers already exist for the amplification and sequencing of the rest of the rDNA. Different regions of the rDNA have different degrees of nucleotide variation. The 5S, 5.8S, 18S and LSU all have sequence dependent functions. This limits the degree of variability between species. The ITS and NTS have functions that are independent of sequence, as far as is currently known. This increases the variability of their sequence between close species. However, it also increases the potential variability of sequence differences amongst strains of the same species. This is the main advantage of the ITS region for species-specific primers. It is also the main disadvantage.

1.9.2 Housekeeping genes

Housekeeping genes perform essential functions within a cell. They are constitutively expressed due to their constant requirement. DNA sequence that is bound to remain within functional constraints cannot be free to evolve unchecked. At first glance, this would seem unsuitable for distinguishing between species as enough variation over a small enough area to be exploited by PCR primers would be unlikely. The functional constraints preventing sequence change make some crucial regions of the proteins identical, such as enzyme active sites and structural elements. However, despite this, the inherent redundancy of the genetic code allows for nucleotide substitutions even between closely related species without altering the protein sequence. A triplet of nucleotides called a codon encodes a single amino acid. The first nucleotide position is essential for the type of amino acid coded for. The second nucleotide is not as essential and for some amino acids, the third position is arbitrary. A nucleotide substitution resulting in no change in the amino acid sequence is referred to as a synonymous substitution. Changes in the amino acid are due to non-synonymous substitutions primarily in the first and second codon positions. Synonymous substitutions between closely related species are common and can be exploited using specific PCR primers.

1.9.3 Species-specific genes

The ideal genetic target for specific primers is one that occurs only in the target species and nothing else. Species-specific genes are relatively common in bacteria with certain pathogenicity factors. They are not as common in yeasts and are more difficult to find due to the larger genomes of eukaryotes and less full genome sequences.

1.9.4 Primer design strategy and parameters

The strategy for designing species-specific primers requires the sequence data of two sibling species. Sibling is determined by D1 / D2 sequence comparisons. The species with the least number of differences to the target is considered to be the sibling. The sequences are aligned and then primers are designed by hand to encompass as many of the nucleotide differences between the two as possible. Additional sequences from other related species should ideally be used in the comparison, especially for closely related genera.

The parameters of primer design are that they should be around 24bp in length, reducing the likelihood that the same sequence is repeated elsewhere in the genome. They should have as similar melting temperatures as possible. All primers should end on the same nucleotide, preferably a G. If this is not possible then they have to end on non-complementary nucleotides. A GC clamp is most preferable. This requires two consecutive bases at the 3' end that form 3 hydrogen bonds (G or C), ensuring that the primer binds tightly to the target at the position where polymerisation is initiated. As many of the mismatches in the non-target species should be clustered at the 3' end as possible. This will ensure that the primer does not bind tightly to the non-target at the 3' end, thus hindering polymerisation. Primers should avoid a consecutive run of more than two of the same base as these may act like an arbitrary primer and anneal to a complement that will undoubtedly be present elsewhere in the genome.

Multiple forward and reverse primers will be designed for each target species according to the above criteria. They will then be tested in all combinations on a selection of target and sibling strains to determine pairs that are potentially species-specific. Promising primers will then be tested on as many strains as possible to verify their specificity.

The ultimate aim is to have a single reaction mixture for a genus. After a PCR at standardised conditions, a single band of predefined size on an agarose gel will indicate the presence of a species.

Chapter 2. Materials and Methods

2.1 Yeast strains and culture conditions

A list of the yeast strains used is shown in tables 2.1 to 2.4. Yeast liquid cultures were grown in 10ml YPD medium (1% yeast extract, 2% glucose, 1% Bactopeptone (all w/v)). Yeast plated cultures were grown on YPD agar (as YPD plus 2% agar (w/v)). *Saccharomyces sp.*, *Schizosaccharomyces sp.* and *Zygosaccharomyces bailii*, *Z. bisporus*, *Z. mellis* and *Z. rouxii* were incubated at 28°C for approximately 20hours, with shaking in liquid culture. *Z.lentus* and *Z.kombuchaensis* were incubated at 25°C, without shaking in liquid culture, for approximately 30hours. *Brettanomyces* and *Dekkera* species were incubated at 28°C, with shaking in liquid culture, for approximately 30hours. This time was increased to 48hours if liquid cultures were inoculated from solid medium, as the yeast requires time to convert from filamentous to budding form.

2.2 DNA preparation

2.2.1 Cell number quantification

Overnight cultures were diluted 10-fold in water. 10µl were placed between a coverslip and haemocytometer grid and left to settle for a few minutes. The numbers of cells in 5 diagonal squares of the haemocytometer were counted and the cell titre calculated.

2.2.2 Clean DNA preparation for PCR

DNA for PCR was prepared using the Qiagen DNeasy Tissue Kit (Qiagen, UK) following the manufacturers guidelines for yeast cultures. 1ml of liquid culture was used at 5 x 10^7 cells/ml.

Species	Number	Local Code	Culture collection codes	Origin	Source
S. bayanus	1	565	NCYC 2669 CBS 7001		Insect: *Mesophylax adopersus*
S. bayanus	2	59	CMCC 2271		Cider apples
S. bayanus	3	5114			
S. bayanus	4*	4235	CBS 1604		Pear pulp
S. bayanus	5	24	CMCC 2653		Carbonated ice tea
S. bayanus	6	5116	VKMY 361		
S. bayanus	7		NCYC 509 CBS 395		Blackcurrant juice
S. cariocanus	1	5251	CBS 7994 IFO 10948	Brazil	Fruit fly, *Drosophila* sp.
S. cariocanus (T)	2	5252	NCYC 2890 CBS 7995 IFO 10947	Brazil	Fruit fly, Drosophila sp.
S. cariocanus	3	5418	YPS 145	USA	Soil
S. cariocanus	4	5413	YPS 152	USA	Soil
S. cariocanus	5	5416	YPS 158	USA	Soil
S. cariocanus	6	2666	YPS 125	USA	Soil
S. cariocanus	7	4259	YPS 138	USA	Soil
S. cariocanus	8	5411	YPS 150	USA	Soil
S. cariocanus	9	5415	YPS 155	USA	Soil
S. cariocanus	10	4241	DBVPG 6304	USA	Fruit fly, Drosophila sp.
S. cariocanus	11	5412	YPS 151	USA	Soil
S. cerevisiae	1	282		Netherlands	Spoiled ice tea
S. cerevisiae	2	244	AS2.378		Soy sauce
S. cerevisiae	3	3101	GB2333		Commercial active dried yeast
S. cerevisiae	4	4238	CBS 5635	South Africa	Grape must
S. cerevisiae	5	62	CMCC 2268		Wine yeast
S. cerevisiae	6	291	D3253		Soft drink spoilage
S. cerevisiae	7*	4237	CBS 4054	Spain	Red wine
S. cerevisiae	8	310			
S. kudriavzevii (T)	1	5247 / 5108	NCYC 2889 IFO 1802	Japan	Partially decayed leaf
S. kudriavzevii	2	5248 / 5107	IFO 1803	Japan	Partially decayed leaf
S. mikatae (T)	1	5106	NCYC 2888 IFO 1815	Japan	Soil
S. mikatae	2	5105	IFO 1816	Japan	Partially decayed leaf
S. paradoxus	1	5103			Ex. Ed Louis
S. paradoxus	2	4240	CBS 5829	Denmark	Soil
S. paradoxus	3	5102			
S. paradoxus (T)	4	4239	NCYC 2600 CBS 432	Russia	Sap exudate of tree
S. paradoxus	5	5101			
S. paradoxus	6	4234	CBS 1515		Juice of Fendant grapes
S. pastorianus (T)	1		NCYC 392		Brewing
S. pastorianus	2	4667			Brewing
S. pastorianus	3	4417			Brewing
S. pastorianus	4	4243	NCYC 369 CBS 1513		Brewers bottom yeast
S. pastorianus	5	4242	NCYC 400 CBS 1260		Brewers top yeast
S. pastorianus	6		NCYC 2801 CBS 1503		Brewer's bottom yeast
S. pastorianus	7		NCYC 453		Brewing

Table 2.1: *Saccharomyces* strains used and their culture collection codes, source and origin, where known. Local strains are available from the University of Bath culture collection.

Species	Number	Local Code	Culture collections	Origin	Source
B. custersianus	1	ISA 1981			
B. custersianus	2	CBS 8347	CBS 8347	Netherlands	Olives
B. naardenensis	1	27	CMCC 2682	UK	Spoiled carbonated ice tea
B. naardenensis	2	ISA 1721			
B. naardenensis	3	129	NCYC 899		Spoiled carbonated soft drink
B. nanus (T)	1	ISA 1945	CBS 1945		
D. anomala		311 (*brux*)			
D. anomala	1	D3288	NCYC D3288	UK	Spoiled beer
D. anomala	2	502			Kombucha drink
D. anomala	3	506			Kombucha drink
D. anomala	4	ISA 1652			
D. anomala	5	ISA 1653			
D. anomala	6	C384		Chile	Wine
D. bruxellensis	1	128	NCYC 823		Lambic beer
D. bruxellensis	2	148		Netherlands	Spoiled ice tea
D. bruxellensis	3	ST1			
D. bruxellensis (T)	4	C9	NCYC 823 CBS 74		
D. bruxellensis	5	306	LP 64	Belgium	Spoiled ice tea
D. bruxellensis	6	ISA 1644			
D. bruxellensis	7	311(*anomala*)		UK	Spoiled beer
D. bruxellensis	8	ISA 1650			
D. bruxellensis	9	ISA 1749			
D. bruxellensis	10	ISA 1717			
D. bruxellensis	11	ISA 1791			
D. bruxellensis	12	NS849		Chile	Wine
D. bruxellensis	13	C390		Argentina	Wine
D. bruxellensis	14	C388		Argentina	Wine
D. bruxellensis	15	C389		Argentina	Wine

Table 2.2: *Brettanomyces / Dekkera* strains used and their culture collection codes, source and origin, where known. Local strains are available from the University of Bath culture collection.

Species	Number	Local Code	Culture collections	Origin	Source
Z. bailii	1	3528	NCYC 1400		
Z. bailii	2	3257			
Z. bailii	3	3135	NCYC 1427		
Z. bailii	4	2422	NCYC 1766		
Z. bailii	5	4351	NCYC 1556		
Z. bailii	6	3532			
Z. bailii	7	3530			
Z. bailii	8	3531			
Z. bailii	9	3469	NCYC 563		
Z. bailii	10	362		Turkey	Factory
Z. bailii (T)	11	5220	NCYC 1416 CBS 680		
Z. bisporus	1	257		Vlaardingen	
Z. bisporus	2	304	NCYC 1496 CBS 1083	Java	Kombucha tea
Z. bisporus (T)	3	5221	NCYC 1495 CBS 702		
Z. bisporus	4	134	NCYC 1555		Salad cream
Z. bisporus	5	104		UK	Tomato based product
Z. bisporus	6	2457		UK	Tomato based product
Z. bisporus	7	390		Turkey	Factory
Z. bisporus	8	3533 (*bailii*)			
Z. kombuchaensis	1	102	NCYC 2969	Russia	Kombucha tea
Z. kombuchaensis	2	199	NRRL Y-27162	USA	Kombucha tea
Z. lentus	1	39	CMCC 3590	France	Wine
Z. lentus	2	158			
Z. lentus	3	161			
Z. lentus	4	37	NCYC 2789 CMCC 3588	UK	Whole orange juice
Z. lentus	5	103	NCYC 2406	UK	Tomato based product
Z. lentus	6	160			
Z. lentus	7	159			
Z. lentus	8	40	CMCC 3589	UK	Spoiled orange beverage
Z. mellis	1	4229 (MS 142)	CBS 738	Canada	Fermenting honey
Z. mellis	2	4226 (MS 139)	CBS 684		Honey
Z. mellis (T)	3	2461	NCYC 2403		
Z. rouxii	1	2452 (MS 115)		UK	Fermenting marmalade
Z. rouxii	2	4219			Fruit concentrate
Z. rouxii	3	4220			Starch hydrolysate
Z. rouxii	4	4218			Fruit concentrate
Z. rouxii	5	3222 (MS 33)		UK	Fermenting drink
Z. rouxii	6	3208 (MS 16)		UK	Fermenting strawberry jam
Z. rouxii	7	3203 (MS 34)		UK	Wine
Z. rouxii	8	5222			
Z. rouxii	9	598		Spain	
Z. rouxii	10	4216			High fructose syrup
Z. rouxii	11	4227 (*mellis*)	CBS 726	Germany	Wine grapes

Table 2.3: *Zygosaccharomyces* strains used and their culture collection codes, source and origin, where known. Local strains are available from the University of Bath culture collection.

Species	Number	Local Code	Culture collections	Origin	Source
L. cidri (T)	1	46	NCYC 1567		
L. fermentati	1	45	NCYC 2508		
T. delbruckeii	1	137	NCYC 582		Strawberries
T. delbruckeii	2	207		Vrumona	Factory
T. microellipsoides (T)	1	42	NCYC 2568		
T. microellipsoides	2	296	NCYC 411		
Zt. florentinus	1	3226	NCYC 2513		
Zt. florentinus	2	44	CMCC 2270		
Zt. florentinus	3	43	NCYC 5213		
Zt. mrakii	1	5222	CBS 4218		
Zt. mrakii	2	4230		Brazil	Pitanga fruit

Table 2.4: Other strains used and their culture collection codes, source and origin, where known.

2.2.3 Rapid cell lysis for PCR

Genomic DNA was prepared for a rapid PCR assessment using *micro*LYSIS® solution (Microzone Limited, UK) following the manufactures guidelines. 1µl high density liquid culture was mixed with 19µl *micro*LYSIS® solution and put through the following thermal cycle:

65°C for 5mins

96°C for 2mins

65°C for 4mins

96°C for 1min

65°C for 1 min

96°C for 30seconds

20°C hold

1µl of this solution was used for subsequent PCR reactions.

2.3 Amplification

2.3.1 Amplification with new primers

The following PCR conditions were used for preliminary amplifications with designed primers, ribosomal primers and universal fungal primers: 0.5 units *Taq* DNA polymerase, 10% PCR buffer (containing 10mM Tris-HCl, 50mM KCl and 0.1% Triton® X-100 per reaction), 1mM dNTP mix, 2.5mM $MgCl_2$ (all Promega, UK), 1mM each primer (Invitrogen, UK) and 2.5µl template DNA. Reactions were made up to 25µl with MilliQ water. The PCR thermal cycler (PTC-100™ Programmable Thermal Controller, MJ Research Inc) conditions were:

94°C for 5mins,

35 successive rounds of

94°C for 1 min

55°C for 1 min

72°C for 1 min

final extension step of 72°C for 2mins

2.3.2 Analysis of amplified PCR products

Routine agarose gel electrophoresis was performed using 1% (w/v) agarose in TBE buffer (0.9M Tris-borate, 2mM EDTA). Gels were run at 100V for approximately 1 hour.

2.4 Optimisation

2.4.1 Temperature gradient PCR

12 duplicate 25µl PCRs were set up as in section 2.3.1 for each temperature gradient assessed. Reactions were performed using a DNA Engine Dyad® Peltier Thermal Cycler (MJ Research Inc) using temperature gradients of 55 to 72°C for *Z. lentus* F2 / R1 and 50 to 70°C for *Z. bailii* F1 / R2 and *Z. kombuchaensis* F3 / R1.

2.4.2 Magnesium ion gradient PCR

Multiple replicate PCR reactions were prepared and performed as in section 2.3.1, but Mg^{2+} ion concentrations were adjusted to molarities of 1.0mM, 1.5mM, 1.6mM, 1.7mM, 1.8mM, 1.9mM, 2.0mM and 2.5mM. Final reaction volume was made up to 25µl with MilliQ water.

2.5 Sequences

2.5.1 Sequences from databases

NCBI accession numbers [108] of sequences used to design primers for sequencing the *Zygosaccharomyces HIS3* gene are listed in table 2.5.

Species	Strain	Gene	Accession number
B. custersianus	CBS 74^T	26S (D1 / D2)	U76199
B. naardenensis	CBS 6042^T	26S (D1 / D2)	U76200
B. nanus	CBS 1945^T	26S (D1 / D2)	U76197
D. anomala	CBS 8139^T	26S (D1 / D2)	U84244
D. bruxellensis	CBS 74^T	26S (D1 / D2)	U45738
L. cidri	NCYC 1567^T	26S (D1 / D2)	U84236
L. fermentati	CBS 707^T	26S (D1 / D2)	U84239
L. thermotolerans	CBS 6340^T	26S (D1 / D2)	U69581
L. waltii	NCYC 2644^T	26S (D1 / D2)	U69582
S. bayanus	CBS 380^T	26S (D1 / D2)	U94391
S. cariocanus	NRRL Y-27337	26S (D1 / D2)	AY046086
S. cerevisiae	NCYC 505^{NT}	26S (D1 / D2)	U44806
S. kudriavzevii	NCYC 2889^T	26S (D1 / D2)	AB040996
S. paradoxus	NCYC 2600^{NT}	26S (D1 / D2)	U68555
T. delbrueckii	CBS 1146^T	26S (D1 / D2)	U72156
T. franciscae	NRRL Y-17532^T	26S (D1 / D2)	U73604
T. globosa	NCYC 820^T	26S (D1 / D2)	U72166
T. microellipsoides	NCYC 2568^T	26S (D1 / D2)	U72160
T. pretoriensis	NRRL Y-17251^T	26S (D1 / D2)	U72157
Z. bailii	NCYC 1416^T	26S (D1 / D2)	U72161
Z. bailii	NCYC 1416^T	ITS1	AJ229067
Z. bailii	NCYC 1416^T	ITS2	AJ229070
Z. bailii	ISA 1307	LEU2	AJ292544
Z. bailii	ATCC 36947	HIS3	AY050224
Z. bailii	NCYC 1416^T	HIS3	AJ634595
Z. bisporus	NCYC 1495^T	26S (D1 / D2)	U72162
Z. bisporus	NCYC 1495^T	ITS1	X84642
Z. bisporus	NCYC 1495^T	ITS2	X84643
Z. kombuchaensis	CBS 8849^T	26S (D1 / D2)	AF399791
Z. kombuchaensis	CBS 8849^T	ITS1	AY046193
Z. kombuchaensis	CBS 8849^T	ITS2	AY046193
Z. lentus	NCYC D2627^T	26S (D1 / D2)	AF339888
Z. lentus	NCYC D2627^T	ITS1	AY046194
Z. lentus	NCYC D2627^T	ITS2	AY046194
Z. mellis	NCYC 2403^T	26S (D1 / D2)	U72164
Z. mellis	NCYC 2403^T	ITS1	Z48349
Z. mellis	NCYC 2403^T	ITS2	Z48364
Z. rouxii	NCYC 568^T	26S (D1 / D2)	U72163
Z. rouxii	NCYC 568^T	ITS1	X84644
Z. rouxii	NCYC 568^T	ITS2	X84645
Z. rouxii	NCYC 568	LEU2	AF314095
Z. rouxii	NCYC 588	HIS3	Y18561
Zt. florentinus	CBS 746^T	26S (D1 / D2)	U72165
Zt. mrakii	CBS 4218^T	26S (D1 / D2)	U72159

Table 2.5: NCBI accession numbers of all downloaded sequences

2.5.2 Sequencing primer design

The primer sequences used to sequence the *Zygosaccharomyces sp. HIS3* gene, *Saccharomyces sp. PRI1*, *MEX67*, *FAL1*, *DBP6*, *SEC31*, *VAS1* and *SPB1* and yeast ribosomal genes [20, 24, 88] are shown in tables 2.6 and 2.7. Primers designed to sequence a gene from all species in a genus were based upon a ClustalW alignment of existing sequences of sibling species. Primers were localised to highly conserved regions between siblings with as few nucleotide differences over a 24 base region as possible. Any nucleotide differences were incorporated into the primer sequence as degenerate positions. All primers were designed to end on a guanine base and had as similar GC content and T_m as possible. Invitrogen, UK and TAGN, UK synthesized all primers.

***Saccharomyces* genes**	**Primer sequence, 5' to 3'**	**Tm (°C)**	**G.C (%)**	**Length (bp)**	**Size of product**
PRI1					
F2	GAA YCA YTC SCC CAA RCC TTC WAG	46 - 58	46 - 58	24	approx 1.1kb
R2	ACG TTC YCK YTT VAY YGA ACC CAG	50 - 59	38 - 63	24	
SPB1					
F2	HAA RGA GAA RGG TTA TCG TGC TCG	52 - 56	42 - 54	24	approx 2 kb
R1	ATA YTT ACC YTT AAC RCC CTT TGG	48 - 55	33 - 46	24	
SEC31					
F3	WRT TTA TGA AGC TCA TTC YGA ATG	49 - 51	29 - 38	24	1736bp
R1	TGT ATT TCT TTT CYT TCA TTA RTG	46 - 48	21 - 29	24	
VAS1					
F3	GAT GAA GAT GTY TTA GAY ACC TGG	52 - 51	38 - 46	24	1312bp
R1	TRG TTT GYT CRA TAC CRY TCT TGG	50 - 58	38 - 50	24	
FAL1					
F2	KYT RAA AGA HGA YTT ACT TCG AGG	47 - 52	29 - 50	24	1031bp
R2	CWG YTY TRG TRA TGA AGT TTA TAG	43 - 48	25 - 42	24	
MEX67					
F2	TYA GYR TRA GRA ACT GGC AGA ATG	51 - 59	33 - 54	24	approx 1.6kb
R2	TTT AYG GCR ACT TCR TAG TTC CAG	54 - 58	38 - 50	24	
DBP6					
F1	CRA GAT TYG AYC CTW SYC ARY TG	42 - 57	35 - 61	23	484bp
R1	TAA TTY STA TTY ARR TTC ATT GG	38 - 46	17 - 35	23	
SEC24					
F1	ACA AGA AAC GTG TTT ACC CAC ARG	52 - 54	42 – 46	24	
F2	ATG ARC TAT GSD CAR ATG GGA ATG	51 - 58	38 – 50	24	
R1	GCV ACT TCC CTA GCA GAT GCA TGG	60 - 62	54 – 58	24	
R2	TGT AAG AAY TCT CTG TAA CTT TCG	52 - 53	33 – 38	24	
R3	CAT TCT YAA ACC TGT GGA ACC ACG	54 - 56	46 – 50	24	
R4	ART CTG ATT GAT CTC TTG GCA TGG	55 - 56	42 – 46	24	
SRP72					
F1	ATT GAA YAT YCA DTT GTC TCA AG	45	26 - 39	23	
F2	AHC CWG CDG AYG TAT TTA GAC G	47 - 53	41 - 55	22	
R1	YCC YTT YTT TGC TAT RTC AAG GG	49 - 51	35 - 52	23	
R2	TTA WTG GHA GCC AYC TTT CAG G	49 - 53	41 - 50	22	

Table 2.6: Primers used to sequence *Saccharomyces* genes

D1 / D2 primers	Primer sequence, 5' to 3'			
NL-1	GCA TAT CAA TAA GCG GAG GAA AAG	54	42	24
NL-4	GGT CCG TGT TTC AAG ACG G	56	58	19
NL-2A	CTT GTT CGC TAT CGG TCT C	53	53	19
NL-3A	GAG ACC GAT AGC GAA CAA G	53	53	19
ITS primers				
ITS 1	TCC GTA GGT GAA CCT GCG G	60	63	19
ITS 2	GCT GCG TTC TTC ATC GAT GC	57	55	20
ITS 3	GCA TCG ATG AAG AAC GCA GC	57	55	20
ITS 4	TCC TCC GCT TAT TGA TAT GC	52	45	20
ITS 5	GGA AGT AAA AGT CGT AAC AAG G	51	41	22
ITS 6	GAA GGT GAA GTC GTA ACA AGG	53	48	21
Universal fungal primers				
NS395F	AGAAACGGCTACCACATCCAAGGAAGG CAGCAGGCGCGCA	73	60	40
NS1654R	CAATCGGTACTAGCGACGGGCGGTGTG TACAAAGGGCAGGGA	72	60	42

Table 2.7: Primers used to sequence rDNA

2.5.3 Amplification for sequencing

Amplification of sequencing material was performed as in section 2.3.1, but reactions were carried out at double quantities in a 50μl final volume and in duplicate. PCR conditions were as follows:

94˚C for 5mins

35 successive rounds of

94˚C for 1 min

55˚C for 1 min

72˚C for 30 seconds per 500bp of anticipated product

final extension step of 72˚C for 2mins

2.5.4 Gel quantification of PCR products

Electrophoresis of quantification gels was performed using 1.5% (w/v) agarose in TBE buffer. Gels were run at 100V for approximately 2 hours for increased resolution. Products were quantified by visual comparison to the MBI Fermentas MassRuler™ Mix, containing bands of known size and quantity.

2.5.5 PCR cleanup

PCR products were cleaned using the Qiaquick PCR Purification Kit (Qiagen UK) according to the manufacturer's guidelines. 20ng/100bp of DNA to be sequenced was purified and eluted in 30µl. Products were then vacuum spun dried for 40mins.

2.5.6 Sequencing

All sequencing was performed by MWG Biotech.

2.5.7 Analysis of sequencing results

Chromatograms of sequence results were assessed to ensure quality data. BioEdit Sequence Alignment Editor version 7.0.2 [109] was then used to align all relevant data from target species of the genus using ClustalW [110].

2.6 D1 / D2 identification of yeast species

Identification of isolates with dubious species assignment was performed using the D1/D2 ribosomal DNA sequence. Isolates were amplified using primers NL-1 and NL-4 described in [24]. PCR reactions were then quantified and prepared for sequencing (section 2.5.4).

The D1/D2 sequence obtained was then aligned with the D1/D2 sequence of the type strains of related species using BioEdit Sequence Alignment Editor version 7.0.2 [109] and species was assigned. If species could not be assigned in this manner, a BLASTn was performed with the sequence compared against the NCBI Fungal database. The GenBank accession numbers for all relevant species type strain D1 / D2 sequences are listed in table 2.5.

2.7 Phylogeny

Phylogenetic trees were generated using MEGA version 3.0 [111].

Chapter 3. *Dekkera / Brettanomyces*

3.1 *Brettanomyces*

A description and history of this target genus is described below with a brief explanation of spoilage characteristics. *Brettanomyces* is the anamorphic genus of *Dekkera.*

3.1.1 Currently accepted species

Brettanomyces custersianus van der Walt (1961)

Brettanomyces naardenensis Kolfschoten & Yarrow (1970)

Brettanomyces nanus (M. Th. Smith, Batenburg-van der Vegte & Scheffers) M. Th. Smith, Boekhout, Kurtzman & O'Donnell (1994)

3.1.2 Description of the genus

Vegetative reproduction – Asexual reproduction is either by multilateral budding or, rarely, by bipolar budding in basipetal succession on a narrow base. Cells are either spheroidal, subglobose to ellipsoidal, frequently ogival, or cylindroidal to elongate. Pseudomycelium is simple or well developed, and branched. One-celled, non-septate mycelium may be formed.

Physiology / biochemistry – Cultures grow slowly and are usually short lived. Acetic acid is produced aerobically from glucose. Fermentation is usually stimulated by molecular oxygen. An extraneous vitamin source is required. Diazonium blue B reaction is negative [112].

3.2 *Dekkera*

A description and history of this target genus is described below with a brief explanation of spoilage characteristics. *Dekkera* is the teleomorphic genus of *Brettanomyces.*

3.2.1 Currently accepted species

Dekkera anomala M. Th. Smith & van Grinsven (1984)

Dekkera bruxellensis van der Walt (1964)

3.2.2 Description of the genus

Vegetative reproduction – Budding cells are spheroidal, subglobose to ellipsoidal, frequently ogival, or cylindroidal to elongate. Pseudomycelium and branched, one-celled, non-septate mycelium are sometimes formed.

Ascospore formation – Asci arise without conjugation, are evanescent, and form 1 – 4 ascospores. Ascospores are hat-shaped or somewhat spheroidal with tangential brims, and tend to agglutinate when released.

Physiology / biochemistry – Cultures are slow growing and generally short lived. Acetic acid is produced aerobically from glucose. Fermentation is usually stimulated by molecular oxygen. An extraneous vitamin source is required. Diazonium blue B reaction is negative [113].

Spoilage characteristics – *Dekkera* and *Brettanomyces* species have similar spoilage properties. They produce sediments, floating particles, acetic and sour-fruity off-flavours. They are predominantly spoilers of low-nutrient soft drinks (cola, tonic, soda water, clear lemon). Fermentation is slow and may not be apparent for weeks. *D. naardenensis* is unable to ferment sucrose. *D. bruxellensis* ferments glucose and sucrose. They are characteristically sensitive to preservatives [107].

Dekkera / *Brettanomyces* species are also very important in wine microbiology. While growth in low numbers produces variable flavour characteristics in wine, overgrowth results in a characteristic "brett" off-flavour. *Brettanomyces* / *Dekkera* tainted wines have been described as "medicinal", "mousy", "horsey" and "wet wool". They have been identified in spoilage from all wine-production areas of the world. The isolation of these yeasts outside of a winemaking environment is very rare. This ecological restriction is thought to be due to their fastidious nutritional requirements [114].

3.3 History of *Dekkera / Brettanomyces*

Van der Walt first described the genus *Dekkera* after observing ascospore formation in *Brettanomyces* species [115]. The genus originally consisted of only *D. bruxellensis* and *D. intermedia*. In the 1980's, a variety of *Brettanomyces* species were seen to produce ascospores resulting in the classification to *Dekkera* of *D. claussenii*, *D. naardenensis*, *D. custersiana*, *D. lambica* and *D. abstinens* [116, 117, 118, 119]. Mitochondrial genome mapping of a variety of *Brettanomyces* and *Dekkera* yeasts concluded that the species *Eeniella nana* was affiliated with *B. custersianus* and *B. naardenensis* [23, 120]. As molecular taxonomy techniques advanced, DNA homology and electrophoretic enzyme profiles established that many of the described species were synonyms and anamorph/teleomorphs of one another. Smith *et al* (1990) [121] described the *Dekkera* species as *Dekkera anomala* (anamorph *B. anomalus*, synonym *B. claussenii*) and *Dekkera bruxellensis* (anamorph *B. bruxellensis*, synonyms *D. intermedia*, *B. abstinens* and *B. custersii*). *B. naardenensis* and *B. custersianus* were distinct *Brettanomyces* species and *Eeniella nana* was too separate to be included in the genus. An rDNA RFLP study established the existence of additional *Dekkera* species, *D. naardenensis* and *D. custersianus* [122] and a 26S rDNA sequencing study established that *Eeniella* should become a *Brettanomyces* synonym [123]. An additional rDNA sequencing study established that *D. custersianus* was sufficiently diverged from the other species to possibly be classified in another genus [124]. It is considered part of the *Dekkera* / *Brettanomyces* genus for this study.

3.4 Results

3.4.1 D1 / D2 sequence alignment

An alignment of the D1 / D2 rDNA regions from each of the type strains in the *Brettanomyces / Dekkera* genus is shown in figure 3.1. A phylogenetic tree constructed from the same alignment is shown in figure 3.2. The relationships illustrated allow determination of closest siblings.

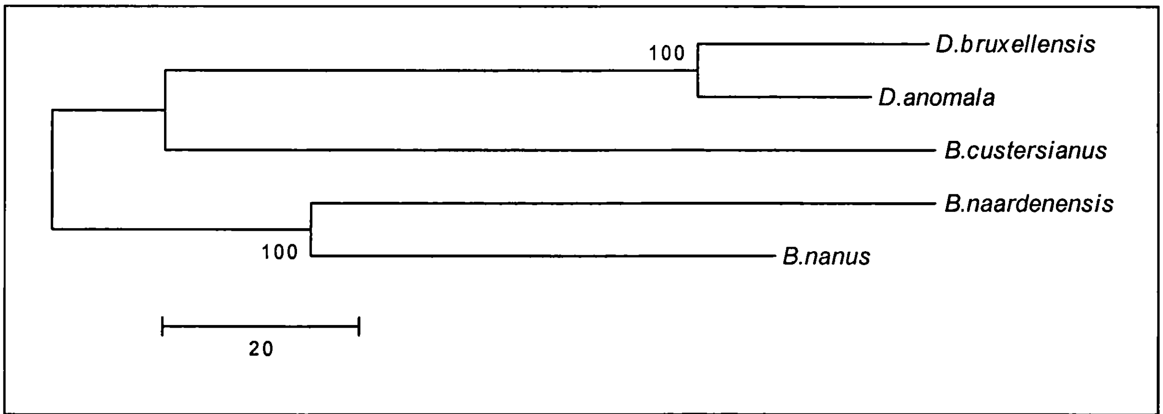

Figure 3.2: Maximum parsimony phylogenetic tree of the D1 / D2 sequences from the *Brettanomyces / Dekkera* type strain of each species. Strains used: *D. bruxellensis* CBS 74, *D. anomala* CBS 8139, *B. nanas* CBS 1945, *B. naardenensis* CBS 6042 and *B. custersianus* CBS 4805 (Number of bootstrap reps = 100).

3.4.2 Specific primer design

A ClustalW alignment was used to compare the D1 / D2 sequences and assess suitable specific primer targets for each species. Specific primer criteria were as follows:

- All primers should end on the same base, or bases not complementary to the end of any other primer.
- Primers should be around 24 nucleotide bases long.
- Primers should have similar GC contents and melting temperatures (T_m).
- Primers should have as many nucleotide mismatches as possible when compared to the closest sibling within the genus.
- Mismatches should be located as close to the 3' end of the primer as possible.

```
                                 10        20        30        40        50        60        70        80        90       100
                         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
B.custersianus_CBS_4805  AAACCAACAGGGATTGCCTCAGTAATGGCGAACGAAGCGGCAACAGCCCAAATTTGAAATCTGGCGTAAGCC---CGAGTTGTAATTTGGAGGTAGACTA
B.nanus_CBS_1945         AAACCAACAGGGATTGCCTCAGTAACGGCGAGTGAAGCGGCAAAAGCTCGAATTTGAAATCAAGTTTCTTG------AATTGTAATTTGAAGAAGCGTCG
B.naardenensis_CBS_6042  AAACCAACAGGGATTGCCTCAGTAACGGCGAGTGAAGCGGCAAAAGCTCGAATTTGAAATCCTCTTCGGAGG-----AGTTGTAATTTGAAGCTGGTTCT
D.bruxellensis_CBS_74    AAACCAACCGGGATTGCCTCAGTAACGGCGAGTGAAGCGGCAAAAGCCCAAATTTGAAATCGGGCAACCGAG--------TTGTAATTTGGAGACGGGACA
D.anomala_CBS_8139       AAACCAACAGGGATTGCCCCAGTAATGGCGAATGAAGCGGCAAGAGCCCAAATTTGAAATCAGGCCCTCGNGGCTTGAGTTGTAATTTGGAGACGGGATA
                                110       120       130       140       150       160       170       180       190       200
                         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
B.custersianus_CBS_4805  TTTGGGTGTCTGTG-----TCAACTAAGTGCCTTGGAATAGGCCGCCGTAGAGGGTGAGAGCCCCGTGAGTTGACTAGCAGTTATTCAGTTTGTCTT---
B.nanus_CBS_1945         CAAGGTGGAGTGAT----AGTG--CAAGTGCCTTGGAACAGGCCGCCAAGGAGGGTGAGAGCCCCGTGA----CACTACCACCGACACCGTCAGCGATGC
B.naardenensis_CBS_6042  TTAGGGAGTTTTGTTTTTGGCTGCGGAAGTGCCTTGGAACAGGCCGCCGTGGAGGGTGAGAGCCCCGTATCGTGGCCAGCAAGATTTCGTATAAGGGCT
D.bruxellensis_CBS_74    CTAGAGAGGAGGAA----GGCGATTAAGTGCCTTGGAACAGGCTGCCGTAGAGGGTGAGAGCCCCGTGAA-TCGCTGGAGACCGATCAATTAGTGCC-C
D.anomala_CBS_8139       CTAGAGAGAGGGAG----GGCGACTAAGTGCCTTGGAACAGGCTGCCGTAGAGGGTGAGAGCCCCGTGAG-TCGCGTGAACTCGATCAATTAGTGCC-C
                                210       220       230       240       250       260       270       280       290       300
                         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
B.custersianus_CBS_4805  -CCGACGAGTCGAGTTGTTTGGGATTGCAGCTCTAAGTGGGTGGTATATTCCATCTAAGGCTAAATAATAGCGAGAGACCGATAGCAAACAAGTACAGTG
B.nanus_CBS_1945         CTCGAAGAGTCGAGTTGTTTGGGAATGCAGCTCTAAGCGGGTGGTATATTCCATCTAAGGCTAAATATTGGCGAGAGACCGATAGCAAACAAGTACAGTG
B.naardenensis_CBS_6042  AGCGAAGAGTCGAGTTGTTTGGGAATGCAGCTCTAAGTGGGTGGTATATTCCATCTAAGGCTAAATATGGGCGAGAGACCGATAGCAAACAAGTACAGTG
D.bruxellensis_CBS_74    GCCGAAGAGTCGAGTTGTTTGGGAATGCAGCTCTAAGTGGGTGGTATATTCCATCTAAGGCTAAATATTAGCGAGAGACCGATAGCAAACAAGTACAGTG
D.anomala_CBS_8139       GCCGAAGAGTCGAGTTGTTTGGGAATGCAGCTCTAAGTGGGTGGTATATTCCATCTAAGGCTAAATATTAGCGAGAGACCGATAGCAAACAAGTACAGTG
                                310       320       330       340       350       360       370       380       390       400
                         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
B.custersianus_CBS_4805  ATGGAAAGATGAAAAGAACTTTGAAAAGAGAGTGAAATAGTACGTGAAATTGTTGAAAGGGAAGGGTATCTGATCCGACGTGGTATTTAGATGTCGCCTG
B.nanus_CBS_1945         ATGGAAAGATGAAAAGAACTTTGAAAAGAGAGTGAAATAGTACGTGAAATTGTTGAAAGGGAAGGGTATTTGATCCGACAGACCTTTTAA----------
B.naardenensis_CBS_6042  ATGGAAAGATGAAAAGAACTTTGAAAAGAGAGTGAAAAAGTACGTGAAATTGTTGAAAGGGAAGGGTATTTGATCCGACATGGTATTTAGATGTCGCTTG
D.bruxellensis_CBS_74    ATGGAAAGATGAAAAGAACTTTGGAAAGAGAGTGAAATAGTACGTGAAATTGTTGAAAGGGAAGGGTATTTGATCCGACATGGTGTTTAGCAGCGGCCCG
D.anomala_CBS_8139       ATGGAAAGATGAAAAGAACTTTGGAAAGAGAGTGAAATAGTACGTGAAATTGTTGAAAGGGAAGGGTATTTGATCCGACATGGTGTTTAGCAGCGGCCCG
                                410       420       430       440       450       460       470       480       490       500
                         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
B.custersianus_CBS_4805  --CCCTTGTGGCGGGTGCTCCATTTTTTTACTGGGCCAGCGTGAGTTAGAAGGGGATGAAAAGGCGGGAACGAATGTAGCCTTCCTCCCTTTGGGAGAGG
B.nanus_CBS_1945         -------------GGTG---------------GTGCCAGCATCGGTTGGCGAAGAAGGATAAAGGTGTGGC-AATGTGCCCCTTCGGGGA-GTTATAGGC
B.naardenensis_CBS_6042  --CCCCCGTGGCGGGCGCTCCATCTTTTTACTGGGCCAGCATCGGTGCTGGGCGGGACAGGAGGTTTCTGTGAATGTAGCCCTTCGGGGAACTTATAGAA
D.bruxellensis_CBS_74    TTCCTCGTGGATGGGTGCACCTGGTTTACACTGGGCCAGCATCGGTTCTGGGAGCCATATACGGGGTTCGTGAATGTGGCCCTTCGATTCTGTCGGAGGG
D.anomala_CBS_8139       TTCCTCGTGGATGGGTGCACCTGGTTTACACTGGGCCAGCATCGGTTCTGGGAGCCATATACGGGGCAGGCGAATGTGGCCCTTCGATTCTGTCGGAGGG
                                510       520       530       540       550       560       570       580       590
                         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|.
B.custersianus_CBS_4805  GGTGTTATAACATTCCTGTTTAGTTCCTTTCTGACTGAGGTCCGCGGTTTTT--TATTACCAAGGACGCTGGCAGAACGAGCAGATACCGC
B.nanus_CBS_1945         CGC----GCACAGACTTCTCTTGGCGACCGAGGACTGCGGAA-GTAGTGTTTACTGCTTCCAAGGATGCTGGCAGAACGAGCAAATACCGC
B.naardenensis_CBS_6042  CAGGAT-TCATACGTCTCTCCCCGGTACCGAGGACTGCGGGA---------------AACCAAGGATGCTGGCATAACGAGCAAATACCGC
D.bruxellensis_CBS_74    TGTTATAGCGCGGACATCTTGTGGCTAGCCGGGACCGGGGACTGCGGTGACT--TGTCACCAAGGATGCTGGCAGAACGAGCAAATACCAC
D.anomala_CBS_8139       TGTTATAGCGCCAGCCAGACGTGGCTATCCGGGACCGGGGACTGCGGCGATTC-TATCGCCAAGGATGCTGGCAGAACGAGCAAATACCGC
```

Figure 3.1: ClustalW alignment of the rDNA region D1 / D2 of the type strains of genus *Dekkera* / *Brettanomyces*. Potential specific primers are highlighted. NCBI accession numbers are listed in table 2.5.

All primers were synthesized by Invitrogen, UK. A list of the potentially specific primers designed for this genus is shown in table 3.1. Amplicon sizes are listed in table 3.2.

Dekkera / Brettanomyces	**Primer sequence, 5' to 3'**	**Tm (°C)**	**G.C (%)**	**Length (bp)**
B.custersianus				
Bcus F1	CCA AAT TTG AAA TCT GGC GTA AG	53	39	23
Bcus F2	AGG TAG ACT ATT TGG GTG TCT GTG	54	46	24
Bcus F3	GTT GAC TAG CAG TTA TTC AGT T	50	36	22
Bcus R1	AAC ACC CCT CTC CCA AAG GGA GG	58	61	23
Bcus R2	CAG AAA GGA ACT AAA CAG GAA TG	51	39	23
B.naardenensis				
Bnaa F1	GGT TCT TTA GGG AGT TTT GTT TTT GG	54	38	26
Bnaa F2	GCA AGA TTT TCG TAT AAG GGC TAG	55	42	24
Bnaa R1	CTA CAT TCA CAG AAA CCT CCT G	52	45	22
Bnaa R2	GTC CTC GGT ACC GGG GAG AGA CG	61	52	25
B.nanus				
Bnan F1	TTT GAA GAA GCG TCG CAA GGT GGA G	62	52	25
Bnan F2	ACT ACC ACC GAC ACC GTC AGC GAT GC	64	62	26
Bnan R1	GAT GCT GGC ACC ACC TTA AAA GG	57	52	23
Bnan R2	CGC CAA GAG AAG TCT GTG CGC GG	64	65	23
D.anomala				
Dano F1	AAA TCA GGC CCT CGN GGC TTG AG	59 - 62	57 - 61	23
Dano F2	GCC CCG TGA GTC GCG TGA ACT CG	64	70	23
Dano R1	TCC CGG ATA GCC ACG TCT GGC TG	62	65	23
Dano R2	CAG CAT CCT TGG CGA TAG AAT CG	59	52	23
D.bruxellensis				
Dbru F1	AAT TTG AAA TCG GGC AAC CG	52	45	20
Dbru F2	GCC CCG TGA ATC GCT GGA GAC CG	63	70	23
Dbru R1	CCC GGC TAG CCA CAA GAT GTC CG	62	65	23

Table 3.1: Sequences of potentially *Brettanomyces* / *Dekkera* species-specific primers

	Expected amplicon sizes for each *Brettanomyces / Dekkera* specific primer pair tested		
Species	**Primer**	R1	R2
B. custersianus	F1	445	472
	F2	405	432
	F3	331	358
B. naardenensis	F1	383	438
	F2	299	354
B. nanus	F1	310	387
	F2	231	308
D. anomala	F1	471	512
	F2	370	406
D. bruxellensis	F1	469	
	F2	369	

Table 3.2: Expected amplicon sizes for all *Brettanomyces / Dekkera* potentially specific primer pairs.

3.4.3 Specific primer testing

All potentially species-specific primers were tested using the generic PCR conditions listed in section 2.3. All reagents, including magnesium ions, primers and dNTPs were in excess and the PCR annealing temperature was relatively low at 55°C. The reaction conditions were like this because the first priority when testing primers was to ensure that all targets were amplified despite any intraspecific polymorphism that might occur within the target loci. The conditions also allowed non-specific reactions to be tailored to become species-specific. Annealing temperature and magnesium ion concentration could be altered to improve the stringency.

Brettanomyces nanus

The D1 / D2 rDNA sequence was used to design two forward and two reverse potentially *B. nanus* specific primers (table 3.1). These allow four combinations of forward and reverse primers to be tested for species specificity. Each primer combination was first tested on the target species and the closest sibling species *B. naardenensis* (figure 3.3). Closest sibling is determined for each species using the target sequence data. The species with the most similar sequence is termed "closest sibling". All primer combinations amplified the target species. The *B. nanus* F2 / R2 combination amplified a non-specific higher molecular weight fragment in the *B. naardenensis* strains (figure 3.3D). This primer pair was thus discarded from subsequent test reactions, as it was not specific.

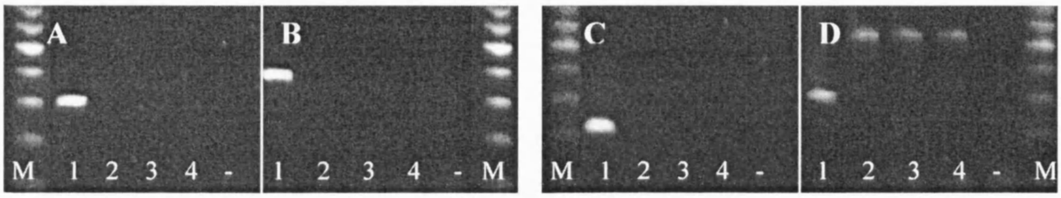

Figure 3.3: Target and closest sibling species amplified with all combinations of forward and reverse *Brettanomyces nanus* primers. (1) *B. nanus* 1, (2) *B. naardenensis* 1, (3) *B. naardenensis* 2, (4) *B. naardenensis* 3. (A) F1 / R1, (B) F1 / R2, (C) F2 / R1, (D) F2 / R2.

The remaining primer combinations were then tested on a selection of *B.custersianus* (figure 3.4), *D. bruxellensis* (figure 3.5) and *D. anomala* (figure 3.6). No non-target species were amplified. The *B. nanus* primer combinations of F1 / R1, F1 / R2 and F2 / R1 are therefore *B. nanus* specific. Refer to table 2.2 for strain labelling.

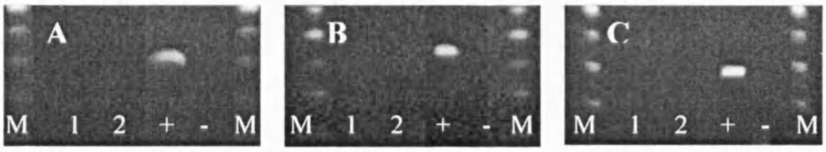

Figure 3.4: *B. custersianus* strains amplified with combinations of the *B. nanus* primers that did not amplify *B. naardenensis*, *D. bruxellensis* or *D. anomala*. (1 – 2) *B. custersianus* 1 – 2. (A) F1 / R1, (B) F1 / R2 (C) F2 / R1

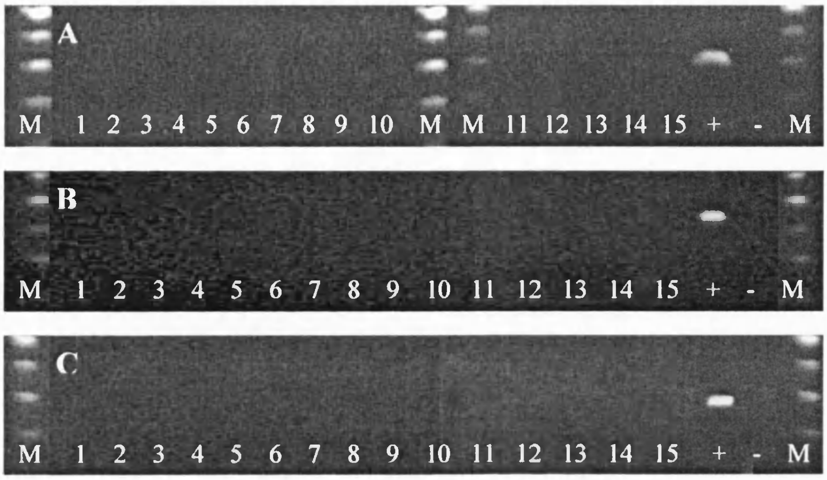

Figure 3.5: *D. bruxellensis* strains amplified with combinations of *B. nanus* primers that did not amplify *B. naardenensis*. (1 – 15) *D. bruxellensis* 1 – 15. (A) F1 / R1, (B) F1 / R2 (C) F2 / R1

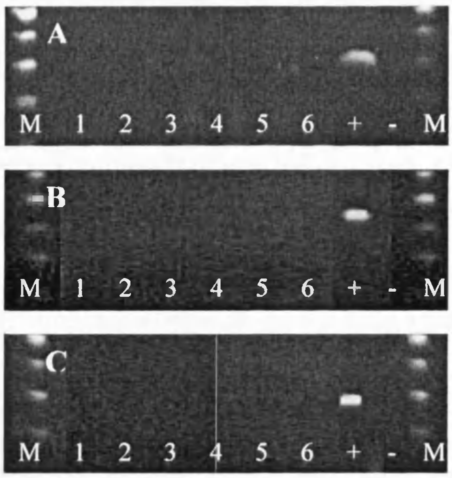

Figure 3.6: *D. anomala* strains amplified with combinations of the *B. nanus* primers that did not amplify *B. naardenensis* or *D. bruxellensis*. (1 – 6) *D. anomala* strains 1 – 6. (A) F1 / R1, (B) F1 / R2 (C) F2 / R1

Brettanomyces naardenesis

The same procedure used for *B. nanus* was followed for *B. naardenensis* specific primers. Two forward and two reverse potentially *B. naardenensis* specific primers were designed using the *B. naardenensis* D1/D2 region as template (table 3.1). Four combinations of forward and reverse primers were then tested for species specificity. The primers were first tested on the three *B. naardenensis* target strains. All were amplified successfully (figure 3.7).

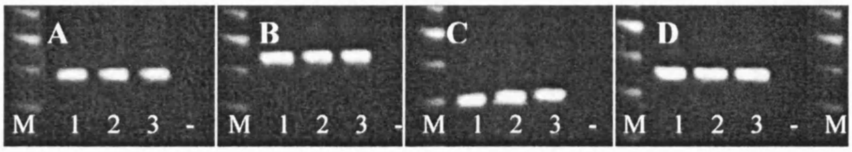

Figure 3.7: Target species amplified with all combinations of forward and reverse *B. naardenensis* primers. (1 – 3) *B. naardenensis* 1 – 3. (A) F1 / R1 (B) F1 / R2 (C) F2 / R1 (D) F2 / R2

All four combinations of primers were then tested on a selection of *B. custersianus* (figure 3.8), *D. bruxellensis* (figure 3.9), *B. nanus* (figure 3.10) and *D. anomala* (figure 3.10) strains. Primer combinations *B. naardenensis* F1 / R2 and F2 / R1 demonstrated weak amplification with *D. bruxellensis* 12 and so were not specific under these conditions. Primer combinations F1 / R1 and F2 / R2 were *B. naardenensis* specific.

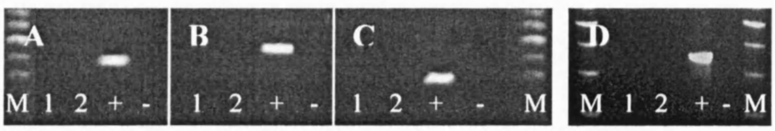

Figure 3.8: *B. custersianus* strains amplified with all combinations of forward and reverse *B. naardenensis* primers. (1 – 2) *B. custersianus* 1 – 2 . (A) F1 / R1 (B) F1 / R2 (C) F2 / R1 (D) F2 / R2

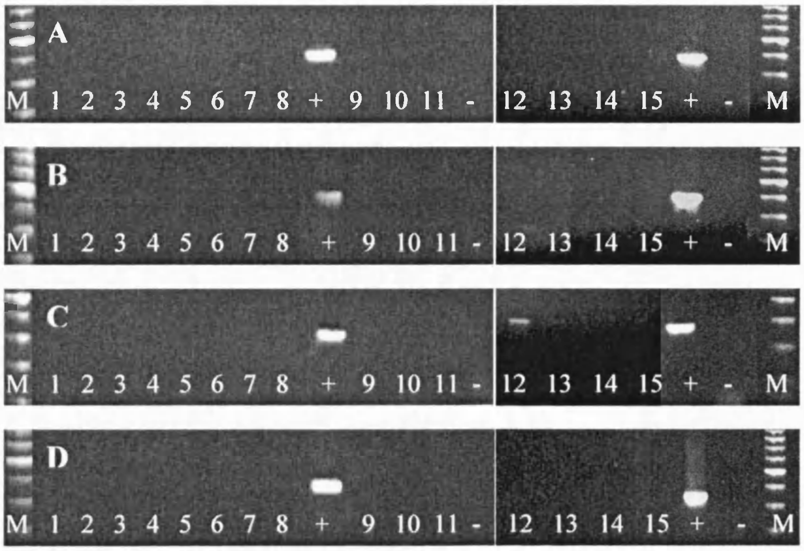

Figure 3.9: *D. bruxellensis* strains amplified with all combinations of forward and reverse *B. naardenensis* primers. (1 – 15) *D. bruxellensis* 1 – 15. (A) F1 / R1 (B) F1 / R2 (C) F2 / R1 (D) F2 / R2

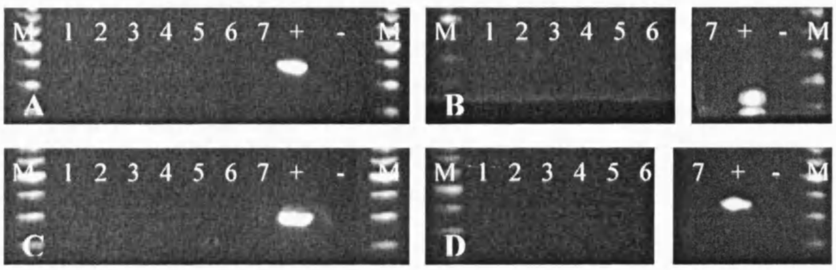

Figure 3.10: *D. anomala* and *B. nanus* strains amplified with all combinations of forward and reverse *B. naardenensis* primers. (1 – 6) *D. anomala* 1 – 6 (7) *B. nanus* 1. (A) F1 / R1 (B) F1 / R2 (C) F2 / R1 (D) F2 / R2

Brettanomyces custersianus

A slightly modified approach was used for *B. custersianus*. Three forward and two reverse primers were designed using the rDNA D1 / D2 region as template (table 3.1). All primer combinations were first tested on a single isolate of each available species (excluding *B. nanus* due to unavailability at the time) (figure 3.11). Three of the primer combinations amplified the closest sibling, *D. anomala* and were abandoned without further investigation. Quickly eliminating the non-specific combinations is a useful approach when there are more numerous primer combinations to begin with.

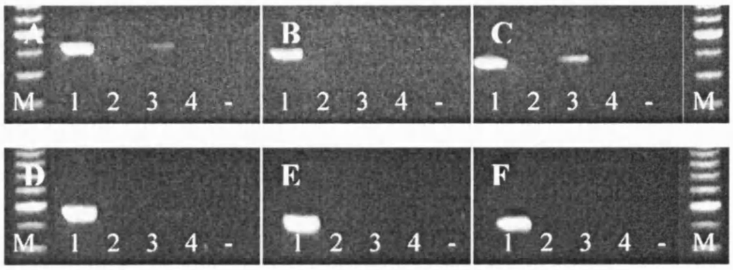

Figure 3.11: Species amplified with all forward and reverse combinations of *B. custersianus* primers. (1) *B. custersianus* 1, (2) *D. bruxellensis* 1, (3) *D. anomala* 1, (4) *B. naardenensis* 1. (A) F1 / R1, (B) F2 / R1, (C) F3 / R1, (D) F1 / R2, (E) F2 / R2, (F) F3 / R2

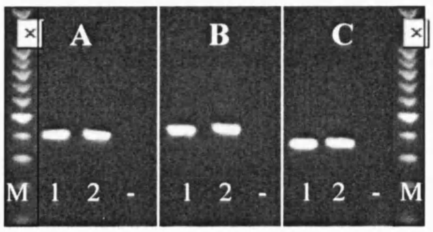

Figure 3.12: *B. custersianus* strains 1 and 2 amplified with all potentially *B. custersianus* specific combinations of forward and reverse primer. (A) F2 / R1, (B) F2 / R2, (C) F3 / R2

The remaining potentially *B. custersianus* specific primers were then tested on target strains (figure 3.12). All target strains were amplified. The same primers were then tested on *D. anomala* (figure 3.13), *D. bruxellensis* (figure 3.14), *B. naardenensis* (figure 3.15) and *B. nanus* (figure 3.16) strains.

Primer combination F2 / R2 amplified a low molecular weight fragment with *B. nanus* (figure 3.16B). Combination F3 / R2 amplified a false positive fragment in *B. nanus* (figure 3.16C). Therefore only primer combination F2 / R1 is *B. custersianus* specific under these reaction conditions.

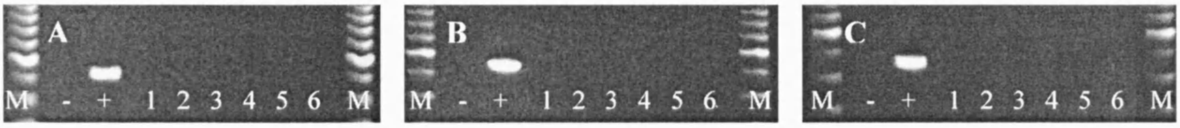

Figure 3.13: *D. anomala* strains amplified with all potentially *B. custersianus* specific combinations of forward and reverse primer. (1 – 6) *D. anomala* 1 – 6. (A) F2 / R1, (B) F2 / R2, (C) F3 / R2

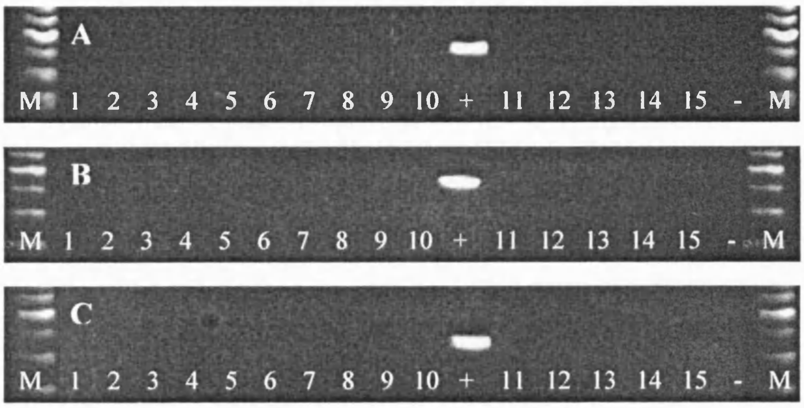

Figure 3.14: *D. bruxellensis* strains amplified with all potentially *B. custersianus* specific combinations of forward and reverse primer. (1 – 15) *D. bruxellensis* 1 – 15. (A) F2 / R1, (B) F2 / R2, (C) F3 / R2

Figure 3.15: *B. naardenensis* strains amplified with all potentially *B. custersianus* specific combinations of forward and reverse primer. (1 – 3) *B. naardenensis* 1 – 3. (A) F2 / R1, (B) F2 / R2, (C) F3 / R2

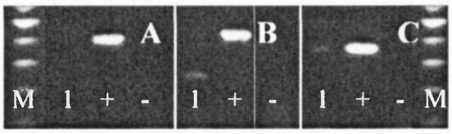

Figure 3.16: *B. nanus* strains amplified with all potentially *B. custersianus* specific combinations of forward and reverse primer. (1) *B. nanus* 1. (A) F2 / R1, (B) F2 / R2, (C) F3 / R2

Dekkera bruxellensis

Only two primer combinations were designed as potential *D. bruxellensis* specific primers: two forward and one reverse primer (table 3.1). The high level of similarity between the D1 / D2 sequences of *D. bruxellensis* and *D. anomala* meant that there were not many nucleotide polymorphisms to direct specific primers towards. The combinations were first tested on the target isolates (figure 3.17). *D. bruxellensis* strains 1 – 8 had previously been confirmed by D1 / D2 sequence analysis. Strains 9 – 11 had been classified only as *Dekkera* species using fatty acid analysis (Virgilio Loureiro, personal communication). Strains 12 - 15 had been identified using physiological methods after being extracted from wine of Chilean and Argentinean origin (Claudio Martinez, personal communication). Using these primers and also negative results from other *Dekkera* and *Brettanomyces* species-specific primers, the *Dekkera* species were confirmed as *D. bruxellensis*. The South American wine strains were shown to contain one *D. anomala* isolate (confirmed by D1 / D2 sequencing), while the rest were *D. bruxellensis*.

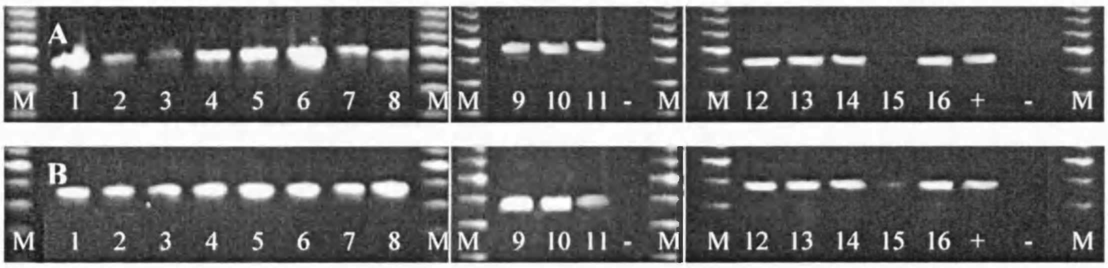

Figure 3.17: Target species amplified with all combinations of forward and reverse *D. bruxellensis* primers. (1 – 14) *D. bruxellensis* 1 – 14, (15) *D. anomala* 6, (16) *D. bruxellensis* 16. (A) F1 / R1 (B) F2 / R1.

The two primer combinations were then tested on the closest sibling *D.anomala* (figure 3.18) and also *B. nanus* (figure 3.19), *B. custersianus* (figure 3.20) and *B. naardenensis* (figure 3.21) strains.

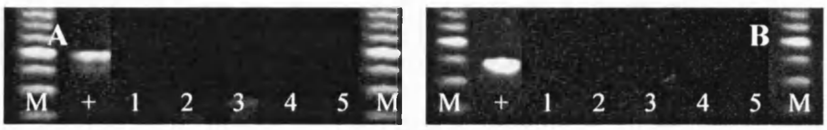

Figure 3.18: *D. anomala* strains amplified with all combinations of *D. bruxellensis* forward and reverse primer. (1 – 5) *D. anomala* 1 – 5. (A) F1 / R1, (B) F2 / R1.

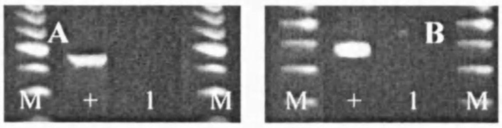

Figure 3.19: *B. nanus* strains amplified with all combinations of *D. bruxellensis* forward and reverse primer. (1) *B. nanus* 1. (A) F1 / R1, (B) F2 / R1.

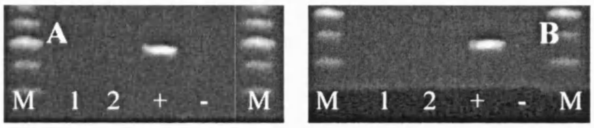

Figure 3.20: *B. custersianus* strains amplified with all combinations of *D. bruxellensis* forward and reverse primer. (1 – 2) *B. custersianus* 1 – 2. (A) F1 / R1, (B) F2 / R1.

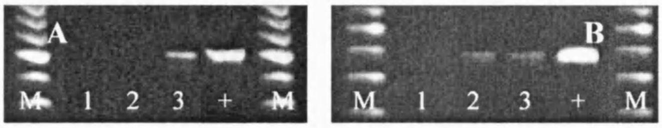

Figure 3.21: *B. naardenensis* strains amplified with all combinations of *D. bruxellensis* forward and reverse primer. (1 – 3) *B. naardenensis* 1 – 3. (A) F1 / R1, (B) F2 / R1.

Figure 3.22: *D. bruxellensis* and *B. naardenensis* strains amplified with all combinations of *D. bruxellensis* forward and reverse primer at 60°C annealing temperature. (1 – 8) *D. bruxellensis* 1 – 8, (9 – 11) *B. naardenensis* 1 – 3. (A) F1 / R1, (B) F2 / R1.

Both primer combinations weakly amplified *B. naardenensis* strains and therefore neither could be deemed *D. bruxellensis* specific under these amplification conditions. The primers were then tested on a selection of target strains and *B. naardenensis* strains at a higher annealing temperature of 60°C (figure 3.22). F2 / R1 was *D. bruxellensis* specific under these conditions.

Dekkera anomala

Two forward and two reverse potentially *D. anomala* specific primers were designed using the rDNA D1/D2 region as template (table 3.1). Four combinations of forward and reverse primers were then tested for species specificity. The primers were first tested on the *D. anomala* target strains (figure 3.23). All were amplified successfully.

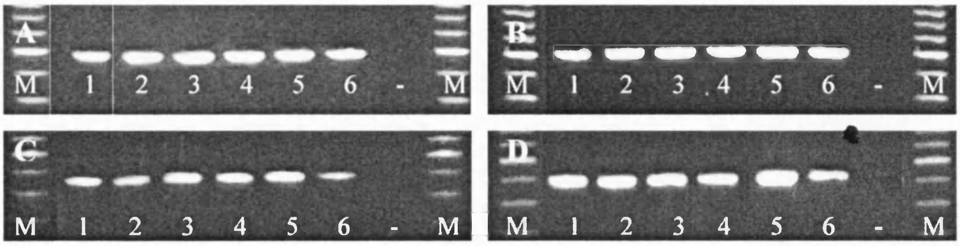

Figure 3.23: Target species amplified with all combinations of forward and reverse *D. anomala* primers. (1 – 6) *D. anomala* 1 – 6. (A) F1 / R1 (B) F1 / R2, (C) F2 / R1, (D) F2 / R2

All four combinations of primers were then tested on *D. bruxellensis* (figure 3.24), *B. custersianus* (figure 3.25), *B. nanus* (figure 3.26) and *B. naardenensis* (figure 3.27) strains.

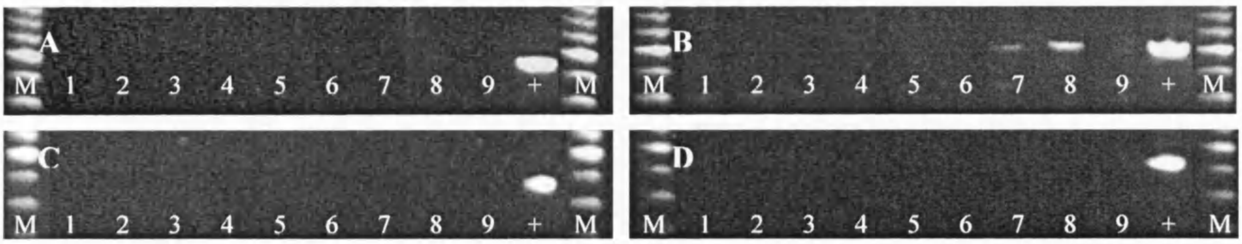

Figure 3.24: *D. bruxellensis* strains amplified with all combinations of *D.anomala* forward and reverse primer. (1 – 9) *D. bruxellensis* 1 – 9. (A) F1 / R1 (B) F1 / R2, (C) F2 / R1, (D) F2 / R2

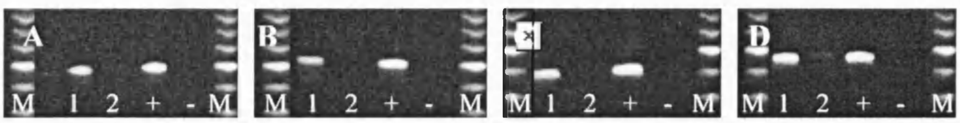

Figure 3.25: *B. custersianus* strains amplified with all combinations of *D.anomala* forward and reverse primer. (1 – 2) *B. custersianus* 1 – 2. (A) F1 / R1 (B) F1 / R2, (C) F2 / R1, (D) F2 / R2

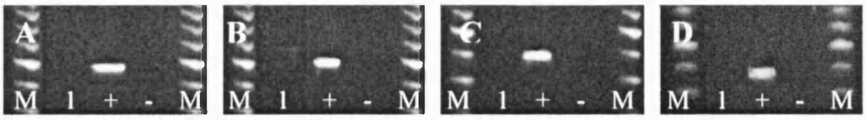

Figure 3.26: *B. nanus* strains amplified with all combinations of *D.anomala* forward and reverse primer. (1) *B. nanus* 1. (A) F1 / R1 (B) F1 / R2, (C) F2 / R1, (D) F2 / R2

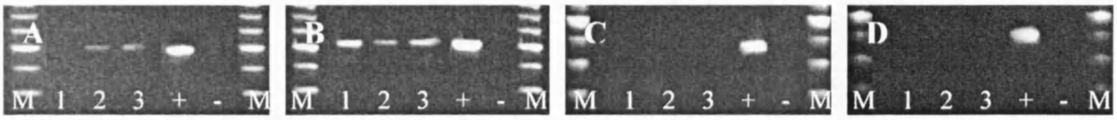

Figure 3.27: *B.naardenensis* strains amplified with all combinations of *D.anomala* forward and reverse primer. (1 – 3) *B. naardenensis* 1 – 3. (A) F1 / R1 (B) F1 / R2, (C) F2 / R1, (D) F2 / R2

Each primer combination strongly amplified *B. custersianus* 1. No primer pairs amplified *B. custersianus* 2. Primer combinations F1 / R1 and F1 / R2 weakly amplified false positive rDNA fragments with all *B. naardenensis* strains, these two combinations were discarded from future work. Primer pairs F2 / R1 and F2 / R2 were tested with all target strains and both *B. custersianus* strains at higher annealing temperatures of 60°C and 65°C (figure 3.28). *B. custersianus* 1 was amplified in all cases despite the higher temperatures.

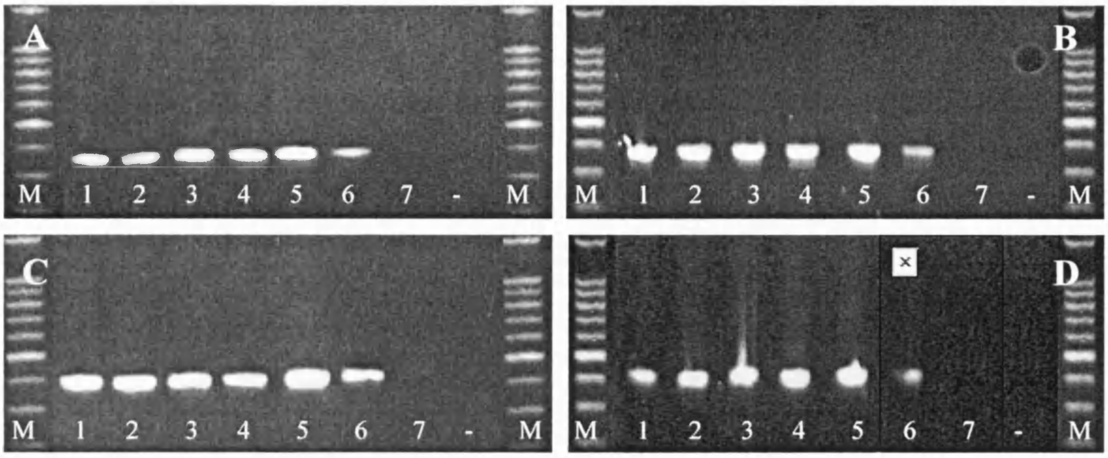

Figure 3.28: *D. anomala* and *B. custersianus* strains amplified with *D. anomala* primers at different annealing temperatures. (1 – 5) *D. anomala* 1 – 5, (6 – 7) *B. custersianus* 1 – 2. (A) F2 / R1 at 60°C, (B) F2 / R1 at 65°C, (C) F2 / R2 at 60°C, (D) F2 / F2 at 65°C.

Both *B. custersianus* 1 and 2 were re-cultured and had fresh genomic DNA prepared. Reactions were identical to previous ones. The strains were then both sequenced in the D1 / D2 region to clarify species identity. Strain 1 (which was amplified) had an identical D1/D2 sequence to the *B. custersianus* type strain. Strain 2 (which was not amplified) had an almost identical D1 / D2 sequence to the type strain (figure 3.29). Attempts to sequence the actual *B. custersianus* 1 amplicon produced with *D. anomala* primers F1 / R2 consistently failed.

```
                                 10        20        30        40        50        60        70        80        90       100
                         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
B.custersianus_T         AAACCAACAGGGATTGCCTCAGTAATGGCGAACGAAGCGGCAACAGCCCAAATTTGAAATCTGGCGTAAGCCCGAGTTGTAATTTGGAGGTAGACTATTT
B.custersianus_1         --------------------------..........................................................................
B.custersianus_2         --------------------------......................................................................G...

                                110       120       130       140       150       160       170       180       190       200
                         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
B.custersianus_T         GGGTGTCTGTGTCAACTAAGTGCCTTGGAATAGGCCGCCGTAGAGGGTGAGAGCCCCGTGAGTTGACTAGCAGTTATTCAGTTTGTCTTCCGACGAGTCG
B.custersianus_1         ....................................................................................................
B.custersianus_2         ..A.AA......TG......................................................................................

                                210       220       230       240       250       260       270       280       290       300
                         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
B.custersianus_T         AGTTGTTTGGGATTGCAGCTCTAAGTGGGTGGTATATTCCATCTAAGGCTAAATAATAGCGAGAGACCGATAGCAAACAAGTACAGTGATGGAAAGATGA
B.custersianus_1         ....................................................................................................
B.custersianus_2         ....................................................................................................

                                310       320       330       340       350       360       370       380       390       400
                         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
B.custersianus_T         AAAGAACTTTGAAAAGAGAGTGAAATAGTACGTGAAATTGTTGAAAGGGAAGGGTATCTGATCCGACGTGGTATTTAGATGTCGCCTGCCCTTGTGGCGG
B.custersianus_1         ....................................................................................................
B.custersianus_2         ....................................................................................................

                                410       420       430       440       450       460       470       480       490       500
                         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
B.custersianus_T         GTGCTCCATTTTTTTACTGGCCAGCGTGAGTTAGAAGGGGATGAAAAGGCGGGAACGAATGTAGCCTTCCTCCCTTTGGGAGAGGGGTGTTATAACATT
B.custersianus_1         ....................................................................................................
B.custersianus_2         ...............................................................................................G....

                                510       520       530       540       550       560       570       580
                         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....
B.custersianus_T         CCTGTTTAGTTCCTTTCTGACTGAGGTCCGCGGTTTTTTATTACCAAGGACGCTGGCAGAACGAGCAGATACCGC---------
B.custersianus_1         ...........................................................................CCGTCTGAA
B.custersianus_2         ..C...................................--......................----------------------
```

Figure 3.29: ClustalW alignment of the rDNA region D1 / D2 of the *B. custersianus* type strain and strains 1 and 2 from this study.

3.5 Discussion

3.5.1 D1 / D2 sequence alignment

D. bruxellensis and *D. anomala* are sibling species and are the most closely related species in the genus. The phylogenetic tree in figure 3.2 visualises the relationship between the species in this genus. The tree topology can be used to assess which species are siblings by tracing back from the tip of a branch to the species it diverges from at the first node. In the case of *D. bruxellensis*, the sibling is *D. anomala*. *B. nanus* and *B. naardenesis* are also siblings. *B. custersianus* is slightly more complex and a ClustalW distance matrix (table 3.3) gives a numerical value to pair wise sequence similarity [110]. *D. anomala* is the closest sibling although *D. bruxellensis* and *B. naardenensis* are almost equally close.

	(1)	(2)	(3)	(4)	(5)
(1) *D. anomala*	0.000	0.066	0.205	0.245	0.230
(2) *D. bruxellensis*	0.066	0.000	0.198	0.238	0.247
(3) *B. nanus*	0.205	0.198	0.000	0.195	0.255
(4) *B. naardenensis*	0.245	0.238	0.195	0.000	0.247
(5) *B. custersianus*	0.230	0.247	0.255	0.247	0.000

Table 3.3: ClustalW distance matrix for the D1/D2 alignment in figure 3.1

Primers targeting the D1 / D2 sequence have the greatest difficulty in distinguishing between sibling species due to a lower degree of sequence variation. Sibling species are the most important when first designing primers to include as many SNP positions as possible.

The taxonomy of this genus is reasonably well established. Not many genes have been sequenced from members of this genus. An NCBI Entrez nucleotide search [125] for sequences derived from "*Dekkera bruxellensis*" gives 53 hits, of which 32 are rDNA sequence, 8 are patents of rDNA sequence, 11 are patents of probes for detecting strains in wine and 2 are for gene sequences (*RAD4* and *COX2*). More sequence data would be required to build a phylogenetic tree with more resolution. However, the D1 / D2 sequences have been shown to give relatively adequate representations of true phylogenies.

3.5.2 Specific primer design

Primer design was a relatively simple task, particularly for *B. nanus* and *B. naardenensis*. There are a large number of polymorphisms within the D1 / D2 sequences of this genus. The most difficult part of primer design in this case was not to find polymorphisms, but to maximise the number of polymorphisms within a primer whilst staying within the design parameters.

3.5.3 Specific primer testing

Brettanomyces nanus

Initially, the most important step in the testing methodology is to be sure that the primers amplify all of the target isolates. In the case of *B. nanus*, only one isolate was available as it is a rare species. It is important to establish that the population of strains available are identical at the primer target loci. In most cases, the number of target strains tested would highlight if a primer target locus is particularly variable. However, with rare species like *B. nanus*, *Z. kombuchaensis* and *S. mikatae* for example, the lack of variability in siblings is taken as proof of a lack of variability in the target species.

The next important step in primer testing is to establish that the primer pairs do not anneal and amplify the same (or any) region in the closest sibling species. The sibling species of *B. nanus* is *B. naardenensis*. The *B. nanus* primer combination F2 / R2 amplified a product of approximately 500bp (200bp larger than the target amplicon) in all *B. naardenensis* strains tested. A BLASTn search [126] against Fungal DNA in the GenBank database revealed no clues as to what this amplicon could be. The advantage of designing multiple forward and reverse primers and testing them in combinations is that results like this do not matter when there are more primer pairs to chose from. The other *B. nanus* primer pairs did not amplify anything when tested against the *B. naardenensis* strains.

Once target amplification and the lack of sibling amplification has been established, the other members of the genus have to be eliminated. No *B. nanus* primer pairs amplified any *B. custersianus*, *D. bruxellensis* or *D. anomala* strains. Primer pairs F1 / R2, F2 / R1 and F1 / R2 were shown to be *B. nanus* specific.

B. custersianus, D. bruxellensis and *B. naardenensis*

B. naardenensis combinations F1 / R1 and F2 / R2 were species specific, the other two pairs weakly amplified *D. bruxellensis* 12 despite there being no significant similarity in BLASTn search of the *D. bruxellensis* D1 / D2 sequence with the primers.

B. custersianus primer pair F2 / R1 was species specific. *B. custersianus* primer combinations F1 / R1, F3 / R1 and F1 / R2 amplified *D. anomala* strain 1. There was no significant similarity between the primer sequence and the *D. anomala* D1 / D2 sequence. Primer F2 / R2 amplified a product approximately 150bp smaller than the target. The source of this amplicon is not apparent from a Fungal BLASTn search. Pair F3 / R2 amplified *B. nanus* despite being the least related species. No significant similarity is seen between the primer sequence and the *B. nanus* D1 / D2 sequence.

All *D. bruxellensis* combinations amplified *B. naardenensis* isolates despite there being no sequence similarity. After increasing the annealing temperature of the reaction, the primer combination F2 / R1 was shown to be species specific.

D. anomala

D. anomala primer pairs were not so straightforward. F1 / R1 and F1 / R2 amplified *B. naardenensis* strains and *B. custersianus* strain 1 (ISA 1981). The remaining *D. anomala* primer pairs F2 / R1 and F2 / R2 were species specific apart from amplifying *B. custersianus* strain 1, even at higher temperatures. Analysis into the nature of this strain found that *B.custersianus* strain 1 has an identical D1 / D2 sequence to the type strain. *B. custersianus* strain 2 (CBS 8347) was not amplified with any of the *D. anomala* primer pairs and has an almost identical D1 / D2 sequence to strain 1 and the type strain on the GenBank database (figure 3.29). A phylogenetic tree illustrating the relationship between the *B. custersianus* strains and the other members of the genus is shown in figure 3.30.

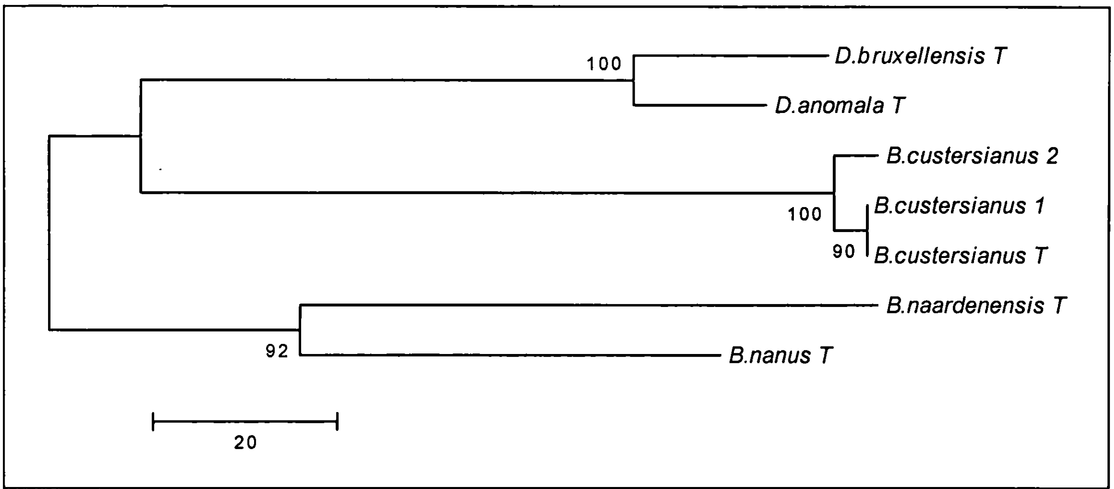

Figure 3.30: Maximum parsimony phylogenetic tree of the D1 / D2 sequences from the *Brettanomyces / Dekkera* type strain of each species and *B. custersianus* strains. Strains used: *D. bruxellensis* CBS 74, *D. anomala* CBS 8139, *B. nanas* CBS 1945, *B. naardenensis* CBS 6042 and *B. custersianus* CBS 4805, 1 and 2 (Number of bootstrap reps = 100).

The *B. custersianus* 1 products with *D. anomala* primers are the same size as the expected *D. anomala* products. It is therefore reasonable to assume that the amplicons are the result of primer annealing at the target loci. Figure 3.31 shows the target sequences in *B. custersianus* strains. There is not a significant degree of similarity to warrant amplification, especially as one strain is amplified and one is not. The primer pairs are more similar in sequence to the *D. bruxellensis* D1 / D2 sequence and yet *D. bruxellensis* strains were only amplified with the *D. anomala* F1 / R2 primer pair under generic PCR conditions. The difference in the behaviour of the B. custersianus strains suggests the possibility that strain 1 contained the rDNA genes of *D. anomala* in the genome. However, the D1 / D2 sequencing reaction performed gave results suggesting the only sequence in the strain was *B. custersianus*. The chromatogram readout was examined for double peaks at positions of polymorphism between *D. anomala* and *B. custersianus*. The reaction was clean with no double peaks to suggest two rDNA sequences were present. A representative portion is shown in figure 3.32.

```
                              10        20
                     ....|....|....|....|...
F1                   AAATCAGGCCCTCGNGGCTTGAG
D.anomala            .......................
D.bruxellensis_T     .....G...AAC....-------
B.custersianus_T     .....T...GTAA..C---C...
B.custersianus_1     .....T...GTAA..C---C...
B.custersianus_2     .....T...GTAA..C---C...

                              10        20
                     ....|....|....|....|....
F2                   GCCCCGTGAG-TCGCGTGAACTCG
D.anomala            ..........-.............
D.bruxellensis_T     .........A-....TG..GAC..
B.custersianus_T     ..........T.GA.TA.C.G.TA
B.custersianus_1     ..........T.GA.TA.C.G.TA
B.custersianus_2     ..........T.GA.TA.C.G.TA

                              10        20
                     ....|....|....|....|...
R1                   TCCCGGATAGCCACGTCTGGCTG
D.anomala            .......................
D.bruxellensis_T     ......C.......AAGAT.TCC
B.custersianus_T     ..AGAA.GGAA.TAAA.A..AAT
B.custersianus_1     ..AGAA.GGAA.TAAA.A..AAT
B.custersianus_2     ..AGAA.GGAA.TAAA.G..AAT

                              10       20
                     ....|....|....|...|....
R2                   CAGCATCCTTGGCGATAGAATCG
D.anomala            .......................
D.bruxellensis_T     ...........T..C. .G..A
B.custersianus_T     ....G.......TA...-..AAA
B.custersianus_1     ....G.......TA...-..AAA
B.custersianus_2     ....G.......TA...-..AAA
```

Figure 3.31: *D. anomala* specific primer target loci in the closest sibling (*D. bruxellensis*) and the *B. custersianus* type strain and strains 1 and 2. Strain 1 is amplified by all of these primers.

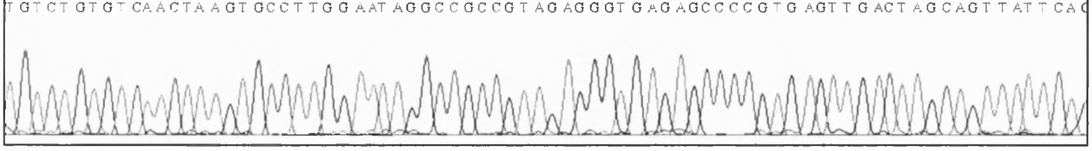

Figure 3.32: A representative portion of the *B. custersianus* strain 1 D1 / D2 sequencing chromatogram readout.

Attempts to sequence the actual *B. custersianus* amplicon that was produced with the *D. anomala* primers were unsuccessful. A clean successful sequencing reaction was impossible to obtain with any of the *D. anomala* primers. The amplicon needs to be ligated into a plasmid and cloned before sequencing can be performed. Alternatively, RFLP analysis of *D. anomala* and *B. custersianus* amplicons would highlight if the amplified *B. custersianus* sequence was identical to that from the target species *D. anomala* or if it was different.

3.6 Conclusion

The heterogeneity within the D1 / D2 sequences of members of the *Brettanomyces / Dekkera* genus is sufficient to exploit with short oligonucleotides to produce species-specific amplicons in a PCR. The nature of *B. custersianus* strain 1 has been difficult to conclusively ascertain and the full nature of this strain is still ambiguous. Some primers amplified species that had a significant number of polymorphisms within the target locus. The copy number of the *Dekkera / Brettanomyces* rDNA genes is unknown. *B. naardenensis* strains were non-specifically amplified most often. This could be due to a higher copy number of rDNA genes in *B. naardenesis*. The amount of target DNA can alter the thermodynamics of a reaction resulting in more non-specific annealing and polymerisation. The generic PCR conditions with all reagents in excess and a purposefully low annealing temperature can facilitate this kind of non-specific amplification. However, these reaction conditions remain necessary to overcome any polymorphism within the target strains to ensure they are amplified. Despite this, successful species-specific PCR primers were designed and tested for *D. bruxellensis*, *B. naardenensis*, *B. nanus* and *B. custersianus*.

Chapter 4. *Zygosaccharomyces*

4.1 Introduction

A description and history of this target genus is described below with a brief explanation of spoilage characteristics.

4.1.1 Currently accepted species

Zygosaccharomyces bailii (Lindner) Guilliermond (1912)
Zygosaccharomyces bisporus H. Naganishi (1917)
Zygosaccharomyces kombuchaensis Kurtzman, Robnett & Basehoar-Powers (2001)
Zygosaccharomyces lentus Steels, Bond, Collins, Roberts, Stratford & James (1999)
Zygosaccharomyces mellis Fabian & Quinet (1928)
Zygosaccharomyces rouxii (Boutroux) Yarrow (1977)

4.1.2 Description of the genus

Vegetative reproduction – Asexual reproduction is by multilateral budding on a narrow base. Cells are spheroidal, ovoidal or elongate. Pseudohyphae, if formed, are generally not well differentiated and true hyphae are not produced.
Ascospore formation - Asci are generally conjugated, with the conjugants frequently presenting a 'dumbbell' configuration. Asci are persistent and produce 1 – 4 ascospores, often equally distributed between the two conjugants.
Physiology / biochemistry – Glucose is fermented, but not galactose. Species variously utilize cadaverine, L-lysine and ethylamine as nitrogen sources, but nitrate is not assimilated. Coenzyme Q-6 is produced. The diazonium blue B reaction is negative [127].

4.1.3 Spoilage characteristics

Z. bailii is renowned for its phenomenal resistance to preservatives and for the damage caused in spoilage of soft drinks, juices, concentrates, jams, wines and ciders. Spoilage is characterised by a sour taste, sediment and slow fermentation up to very high pressure. *Z. bailii* grows relatively slowly and is a fructophile; it ferments fructose in preference to glucose and sucrose very poorly. The natural habitat appears to be fruits, fermenting fruits or mummified fruits where osmotolerance may be an advantage. It grows well in fruit concentrates up to 50% glucose and can grow from extremely low inocula, as few as 1 cell / L. It can also adapt to preservatives if previously exposed to a preservative spillage in the factory and can survive preservative concentrations 4 to 5 times the permitted levels. *Z. bisporus* has the same spoilage profile but is generally less preservative resistant. *Z. microellipsoides* (now *Torulaspora microellipsoides*) has been described as less dangerous than *S. cerevisiae* in the spoilage of soft drinks. It is capable of rapid spoilage of carbonated soft drinks and ferments sucrose and glucose and causes pressure increases and particulates. It is widely distributed in soft drinks plants and moderately preservative resistant. *Z. rouxii* is the most significant spoiler of concentrates, preserved or otherwise due to extreme osmotolerance and a high degree of preservative resistance. It preferentially ferments fructose, cannot ferment sucrose and is capable of growth from very low contamination levels. *Z.lentus* has been isolated from spoiled orange juice, drinks, wine and tomato ketchup. It is fermentative and has a preservative resistance comparable to *Z. bailii*. It is osmotolerant to 60% sugar and can grow at 4°C, which means it has the potential for spoilage of chilled foods [107].

4.1.4 History

Like many species of the *Saccharomyces* complex, these yeasts have previously been classified in the genus *Saccharomyces*, only to be removed and reinstated a varying number of times. Genus *Zygosaccharomyces* Barker was previously a polyphyletic group of species with a core monophyletic *Zygosaccharomyces* "*sensu stricto*" group and a polyphyletic group of *Zygosaccharomyces* "*sensu lato*". The *sensu stricto* species are *Z. bailii*, *Z. bisporus*, *Z. rouxii*, *Z. mellis* and the newest two species *Z. lentus* [128, 129] and *Z. kombuchaensis* [130, 131]. The *sensu lato* species were *Z. microellipsoides*, *Z. fermentati*, *Z. florentinus*, *Z. mrakii* and *Z. cidri*. These were interspersed with species of genus *Torulaspora* Lindner [132]. These species formed the *Zygosaccharomyces* / *Torulaspora* Superfamily of spoilage yeasts. They exhibited similar physiological traits to differing degrees. These traits are fermentation of sugars with a preference for fructose, preservative resistance and osmotolerance. Ribosomal DNA studies revealed close relationships between the species of the two genera [28]. In 2003, Kurtzman [133] resolved the *Saccharomyces* complex yeasts into logical genera based on accumulated research since the establishment of rDNA sequencing methods for yeast taxonomy. The *Zygosaccharomyces sensu lato* species were rearranged into different genera, leaving the *sensu stricto* species as a defined monophyletic *Zygosaccharomyces* genus. *Z. florentinus* and *Z. mrakii* were reclassified as members of the new genus *Zygotorulaspora* while *Z. microellipsoides* was merged into genus *Torulaspora* Lindner. *Z. cidri* and *Z. fermentati* were reclassified with *Kloeckera thermotolerans* and *K. waltii* into the new genus *Lachancea*.

4.2 Results

4.2.1 rDNA sequences

ClustalW alignment of the *Zygosaccharomyces sp.* type strains D1 / D2 region of the large subunit rRNA gene (figure 4.1) shows that this genus is a closely related monophyletic group. The D1 / D2 sequence is too similar to be suitable for the design of successful species-specific primers for any of the 6 members of the genus. Not a lot of alternative sequence data was readily available for any species of the genus, particularly *Z. lentus* and *Z. kombuchaensis*, which are relatively recently discovered species [128, 130]. The next logical step was therefore to examine the rest of the rDNA region.

The 5.8S and 18S sequences were more highly conserved than the D1 / D2 region (data not shown). At first appraisal, the ITS regions seem highly suitable to target with species-specific primers (figures 4.2 and 4.3). However, when each ITS is aligned from strains of the same species, for example *Z. bailii* (figures 4.4 and 4.5), we see that there is intraspecific diversity. For this reason, the ITS sequences were concluded to be unsuitable for this method of species-specific identification.

4.2.2 Housekeeping gene sequences

At the time this project started, the only housekeeping gene sequence data available on GenBank [125] for more than one *Zygosaccharomyces* species was for the genes *LEU2* and *HIS3*. Sequences were available for both *Z. bailii* and *Z. rouxii* and alignments are shown in figures 4.6 and 4.7 respectively. *HIS3* was arbitrarily chosen to sequence for the remaining four members of the genus. Degenerate PCR primers were designed using the *Z. rouxii* and *Z. bailii* sequences as template (all primers are listed in table 4.1 with amplicon sizes in table 4.3). The central part of the gene was first sequenced for all species and then additional primers were used to sequence the 5' and 3' ends of the gene. Figure 4.8 illustrates the method used. The partial *HIS3* gene sequences were then aligned using ClustalW (figure 4.9). The HIS3 genes from multiple *Z. bailii* strains were amplified and sequenced using the primers listed in table 4.2 (amplicon sizes in table 4.3). A ClustalW alignment of these sequences is shown in figure 4.10.

```
                                  10        20        30        40        50        60        70        80        90       100
                          ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.rouxii_NCYC_568         AAACCAACCGGGATTGCCTTAGTAACGGCGAGTGAAGCGGCAAAAGCTCAAATTTGAAATCTGGTACCTTTCGGTGCCCGAGTTGTAATTTGGAGAAAGT
Z.mellis_NCYC_2403        ...........................................G........................C...........................G..C
Z.bisporus_NCYC_1495      .................................................................... ..T.....T..............T....G.C
Z.lentus_NCYC_D2627       ...................C................................................A.......................T....G.C
Z.kombuchaensis_CBS_8849  ....................................................................A.......................T....G.C
Z.bailii_NCYC_1416        .................................................................... .......................T....G.C

                                 110       120       130       140       150       160       170       180       190       200
                          ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.rouxii_NCYC_568         GATTCT GGGACTGGC CCTTGCCTATGTTCCTTGGAACAGGACGTCATAGAGGGTGAGAACCCCGTGAGGCGAGGTGA TCCAGTTCTTTGTAGAACGC
Z.mellis_NCYC_2403        ...... ...G..... G.........................................................A..T A...............G...
Z.bisporus_NCYC_1495      ..C... ...G....T .....T..........................G...........T......T....... AT C...............GT..
Z.lentus_NCYC_D2627       ..C... ...G....T .....T......................................T......T....... AT A...............GT..
Z.kombuchaensis_CBS_8849  ..C... ...G....GA.....T......................................T......T....... CT C...............GT..
Z.bailii_NCYC_1416        ..C... ...G....T .....T..........................G...........T.....AT....... AT C...............GT..

                                 210       220       230       240       250       260       270       280       290       300
                          ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.rouxii_NCYC_568         TTTCGAAGAGTCGAGTTGTTTGGGAATGCAGCTCTAAGTGGGTGGTAAATTCCATCTAAAGCTAAATACAGGCGAGAGACCGATAGCGAACAAGTACAGT
Z.mellis_NCYC_2403        .C....................................A.............................................................
Z.bisporus_NCYC_1495      C...................................................................TT..............................
Z.lentus_NCYC_D2627       ....................................................................TT..............................
Z.kombuchaensis_CBS_8849  ....................................................................TT..............................
Z.bailii_NCYC_1416        C...................................................................TT..............................

                                 310       320       330       340       350       360       370       380       390       400
                          ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.rouxii_NCYC_568         GATGGAAAGATGAAAAGAACTTTGAAAAGAGAGTGAAAAAGGACGTGAAATTGTTGAAAGGGAAGGGCATTTGATCAGACATGGTGTTTTGTGCCCCTCG
Z.mellis_NCYC_2403        ....................................................................................................
Z.bisporus_NCYC_1495      .........................................T..........................................................
Z.lentus_NCYC_D2627       .........................................T.................................................C......TA
Z.kombuchaensis_CBS_8849  .........................................T.................................................C......TA
Z.bailii_NCYC_1416        .........................................T.................................................C........

                                 410       420       430       440       450       460       470       480       490       500
                          ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.rouxii_NCYC_568         CTCCTCGTGGGTGGGGGAATCTCGCAGCTCACTGGGCCAGCATCAGTTTTGGCGGCAGGATAAATCTCTGGGAATGTGGCTTCT  TTCTTCGGGAGGG
Z.mellis_NCYC_2403        ....................................................T.......G...G.CT.............CT.GCC.T..TT..C...
Z.bisporus_NCYC_1495      ..................................................................C..........A...CTGGCA.T. AT.TT..A
Z.lentus_NCYC_D2627       .............A...............................................................A...CTGCT CT.CG...T..A
Z.kombuchaensis_CBS_8849  .............A...............................................................A...CTGCT CT.CG...T..A
Z.bailii_NCYC_1416        .CT........................T......................................C..........A...CTACCACT.CGT..C..AC

                                 510       520       530       540       550       560       570       580       590
                          ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|.
Z.rouxii_NCYC_568         AGTGTTATAGCCC AGGGGAATACTGCCAGCTGGGACTGAGGTATGCGACATT   TTGTCAAGGATGTTGGCATAATGGTTATATGCCGC
Z.mellis_NCYC_2403        G............G.................C..................TC.CG .A.................................
Z.bisporus_NCYC_1495      GAAC.......T. .................C..........A.......T..T   A..........C......................
Z.lentus_NCYC_D2627       GACT......... .................C..........A.......T..T   A..........C......................
Z.kombuchaensis_CBS_8849  GAGT......... .................C..........A.......T..TTT.A..........C......................
Z.bailii_NCYC_1416        GAAC......T.. ............................A.......T..T   A..........C......................
```

Figure 4.1: ClustalW alignment of the *Zygosaccharomyces* species type strains rDNA region D1 / D2. NCBI accession numbers listed in table 2.5

```
                           10        20        30        40        50        60        70        80        90       100
                   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.rouxii_NCYC_568  TAGAAAA------TGAC--GTGAACTCTTAACGGAGTTCTCTC------AAAGTGTTGGAGGGG-AAGGCCTGCGCTTAATTGCGCGGCT----------
Z.mellis_3         C...G..TTGAAA.A.TTC.GAG.GCGGGG.TAT.T.A.NTGGGGGGGA.G....GG.TGAAT.AACCT..........G....CG.CTCT---------
Z.bisporus_3       .......-AATC...TTG.CA.GTCAA.GGG.C...AGAG.TTTTCGGCTCT.TC.TTT.CTTT-TG...........G........TCTAGAGTAGA
Z.lentus_1         .......T-AATC...TTGAA...G..-.---.AG.GAGCA.ACCAAAGGG.C.....TCTCT.C-CG...........G........TCTAGTGGGGG
Z.kombuchaensis_1  .......T-AATC...TTGAAAG.G.T-.TGAA.G.GAGCA.ACCAAAGGG.T.A...TCTCT.----...........G........TCTAGCGGGAG
Z.bailii_2         .......--TTTC...TTGAC..GTC--.GGA.T....G.G.T--------CTCTC..TTTTTTC-.....................TCTAGAGCGGA

                          110       120       130       140       150       160       170       180       190       200
                   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.rouxii_NCYC_568  ----GTTTTT-AATCTCCTCCGCCTT---TGAT--ACACACATTGGAGTTTCT--ACT-TTTTTGTTCTCTTTGGGA---------GGGT-----TCT
Z.mellis_3         ----C...CCTCCTC..G.C.......TGA.T-.AT.......G..........--G..T..C..TA.T.TGA...CTT----TGGG....GTG---.GA
Z.bisporus_3       GGGA.C..----CTA..G..T..G...--ACA..TT.......G..........---...A.....C....T.AG.A.G---ATGAT.AAAA-----...
Z.lentus_1         AGTG...C.---GC...T-.T.AG...TTACA..TA.......G..........---...T..C..C....T.GG.ATT----------...C----TC.G
Z.kombuchaensis_1  AGTA..CC.---GC...T..T.AG...TTACA..TA.......G...........TA...T..C..C....T.GG.AC.TTGTGGGGG...CATTCGTCTC
Z.bailii_2         GGGA...AAACAT.GT.A...T.G...--TCA..TT.......G..........---...T.....A....TC..T..G---AGGAT....------..G

                          210       220       230       240       250       260       270       280       290
                   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.rouxii_NCYC_568  GCTC-TCCCAGAGGTAAA---CACAAACAATCTTTTATTATACTATTA-----ACACAGTCAA--A--TGAATTTTAAAAACAAAAT-ATTCA
Z.mellis_3         ..C.---.............CAA.........GT....T..T.TA.TC..TATATGCACT.....GA.ACAAG.A..CGG.G.A....CG.A.T.
Z.bisporus_3       ATCTAC............---..........A....TA.T.TT....TTAC---.........-.--....A.AC.....A.....---T.
Z.lentus_1         ..C.TG............---..........T....TA.T.TA.-..-----T.A.A......--.CCAA.-AAAC...........-.....
Z.kombuchaensis_1  C...GG...........G---..........T....TA.TATT.-...ATCATT.A.A......--.ACAA.CAAAC...........-.....
Z.bailii_2         T.C.AC.............AA-..........T....T..T.TT.T..TAATATT..........--.AAC..-A.ACT....A..-------..
```

Figure 4.2: ClustalW alignment of the *Zygosaccharomyces* species type strains rDNA region ITS1. NCBI accession numbers listed in table 2.5

```
                          10        20        30        40        50        60        70        80        90       100
                  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.rouxii_NCYC_568 CCTTCTCAAACTTTAC GTTTGGTAGTGAGCGATACTCTACTCTGGAGTTTGCTTGAAAATGGGAGGCCATAGGCGAACGATTGCTTTCCAATCCTGCGG
Z.mellis_3        ..C........GC.TG .C...................TT.TGT.T..A.CTC...GG. .TTT.CTTGA.A.T.GGA.GCCATAGCGGG.G.TA..TT
Z.bisporus_3      ...........AC.TGT.............T........GT-TAA--AAC.-------GG-.TT..CTTGA.ATT.C.A.CC.TT.GGGG.TG.---.TCT
Z.lentus_1        ...........A..CGT......G.....T........GTGTAAC..C..-------G.-.TT..CTTGA.ATT.T.A.CC.TT.GGGTG--.--.TT.
Z.kombuchaensis_1 ...........A.CCGT......G.....T........GTGTAAC..C..-------G.-.TT..CTTGA.ATT.T.A.CC.TT.GGGTG--.--.TT.
Z.bailii_NCYC_1416...........A..CGT.............T........GT T.TTCA...-------GG .TT.ACTTGA.ATT.C.A.CC.TT.GGGG. A.--...T

                         110       120       130       140       150       160       170       180       190       200
                  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.rouxii_NCYC_568 CCCTCTGCTTACTTCCCCTTGTGGGTTGTGGCAGGGGAAAGCGGGAGG CGCCTTGCCACGATA-----GTCGTATTAGGTTTTACCGACTC GGCGAAA
Z.mellis_3        GAGCG..TGGTGGAGGAAC....CCGA.A..TGCA..GCC...CTGCAA.....G.......A.ACGAA.............................G
Z.bisporus_3      AGAGAA.AG.T-------..AG.C.GAAAC.TCT..CTCCT.CTCTATTTT....T-A..C.A.T---......................-...AG.C
Z.lentus_1        G.AAG ATG.T-------...G.C.GA.AC..CCA..C.GC..ATT---TT.T..---..C.A.T---......................-...AG.C
Z.kombuchaensis_1 G.AAG..T-.T-------...G.C.GA.AC..CCA..C.GC..ATT---TT.T.----..C.A.T---......................-...AG.C
Z.bailii_NCYC_1416GTGGG..AG.T-------..AG.C.GAAAC.TCTT.CTC T.CTCT---TT...---A..C.A.T---.......................C.A.AG.C

                         210       220       230       240       250       260       270       280       290
                  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.rouxii_NCYC_568 GTG----AAGAGGTTTGCTTTTTAAAAAGAA-----------------------GCAGGCAGCGTCTGGCTTGACAAAATTCTCAAAGT
Z.mellis_3        .AA-------.TA..GGA.GGGGGG......GCGGAGCTCTCTTTTTTGTTTTGTTTGTTG.TGA.AC .ACGAG...G.GCAG..G..C
Z.bisporus_3      AG.TTGCTGGAGAC.GA.G.GGG.G.T.GA..TATCGAACT-------------TTTCT..GCGCCT.T.GCAA....T.C....T...C
Z.lentus_1        .GT-------.GA.CGGA..G.G.T C.T---------TT--------------TC.TT.TGCGCCT.-.GCAA....TG........C
Z.kombuchaensis_1 .GT-------.GA..GGA..G.G.G.C.T---------CT--------------TC.TT.TGCGCC..-.GCAA....TG........C
Z.bailii_NCYC_1416.G.--ACT.G...A..G.G.GAG.G.T.GC..TATCGAGCTCTGCCTAATTTTTTTTTT..GCGCCT...GCAA....T.C........
```

Figure 4.3: ClustalW alignment of the *Zygosaccharomyces* species type strains rDNA region ITS2. NCBI accession numbers listed in table 2.5

```
                             10        20        30        40        50        60        70        80        90       100
                     ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii_NCYC_1416   GATTGACGAGTCTGGAGTAGTTGTGTTCTCTCTGTTTTTTC-AAGGCCTGCGCTTAATTGCGCGGTCTAGAGCGGAGGGAGTT-----AAACATAGTTAC
Z.bailii_1           .........................................-.-.......................................-----..G.........
Z.bailii_2           C-----------.AA..C.T.........T......CA.---.........................................TTAA-..G..C.A..GT
Z.bailii_3           .................C...........T..........TT.........................................-----..G.......G.
Z.bailii_4           .................C...........T..........TT.........................................-----..G.......G.
Z.bailii_5           .................C...........T..........TT.........................................-----..G.......G.
Z.bailii_8           .................C...........T..........TT.........................................-----..G.......G.
Z.bailii_9           .........................................-.........................................-----............
Z.bailii_13          .........................................-.-.......................................-----..G.........

                            110       120       130       140       150       160       170       180       190       200
                     ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii_NCYC_1416   TCTGGCTTTCAATTTACACACAGTGGAGTTTCTACTTTTTTTATTCTTCTTTGGGAGGATGGGTTCGTCCCACTCCCAGAGGTAAAAACACAAACAATTT
Z.bailii_1           ....................................................................................................
Z.bailii_2           .T....................................................................T...............--............
Z.bailii_3           .T.....................................................................G..T...........--............
Z.bailii_4           .T.....................................................................G..............--............
Z.bailii_5           .T.....................................................................G..............--............
Z.bailii_8           .T.....................................................................G..............--............
Z.bailii_9           ....................................................................................................
Z.bailii_13          ....................................................................................................

                            210
                     ....|....|....
Z.bailii_NCYC_1416   TTTT-TTTTTTTTT
Z.bailii_1           ....T.........
Z.bailii_2           ....T.......CC
Z.bailii_3           ....TA....A...
Z.bailii_4           ....TA....A...
Z.bailii_5           ....TA....A...
Z.bailii_8           ....TA....A...
Z.bailii_9           ....T........A
Z.bailii_13          ....T.........
```

Figure 4.4: ClustalW alignment of the *Zygosaccharomyces bailii* strains rDNA region ITS1.

```
                              10        20        30        40        50        60        70        80        90       100
                     ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii_NCYC_1416   CCTTCTCAAACATTCGTGTTTGGTAGTGAGTGATACTCTGTTTT TCATTTGGGTTAACTTGAAATTGCAAGCCTTTTGGGGAACGCGTGTGGGT GAGT
Z.bailii_1           ............................................A.T.................................... ........... ....
Z.bailii_2           ...............A............................ A.AT....................................T......C...A....
Z.bailii_3           ............................................A.T.................................... ........... ....
Z.bailii_4           ............................................A. .................................... ........... ....
Z.bailii_5           ............................................A. .................................... ........... ....
Z.bailii_8           ............................................A. .................................... ........... ....
Z.bailii_9           ............................................  ..................................... ........... ....
Z.bailii_10          ............................................A.T.................................... ........... ....
Z.bailii_12          ............................................A.T.................................... ........... ....
Z.bailii_13          ............................................ ....................................... ........... ....

                             110       120       130       140       150       160       170       180       190       200
                     ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii_NCYC_1416   TTTAGGCGGAAACGTCTTGCTCTCCTCTTTCCTAACC2AAATGTCGTATTAGGTTTTACCGACTCCGACAGACGGGACTAGGAGATTGGGTGAGTGATAGC
Z.bailii_1           ....................................................................................................
Z.bailii_2           ..........................T.........................................................................
Z.bailii_3           ....................................................................................................
Z.bailii_4           ....................................................................................................
Z.bailii_5           ....................................................................................................
Z.bailii_8           ....................................................................................................
Z.bailii_9           ....................................................................................................
Z.bailii_10          ....................................................................................................
Z.bailii_12          ....................................................................................................
Z.bailii_13          ....................................................................................................

                             210       220       230       240       250       260       270       280
                     ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii_NCYC_1416   AATATCGAGCTCTGCCTAA-------------------TTTTTTTTTT GCGCGCCTTGGGCAAACAATACTCTCAAAGT
Z.bailii_1           ...................-------------------.......... ...............................
Z.bailii_2           ...............T...AAGAAAAAAGATCTCTTTT..........T...............................
Z.bailii_3           ...................-------------------.......... ...............................
Z.bailii_4           ...................-------------------..........T...............................
Z.bailii_5           ...................-------------------..........T...............................
Z.bailii_8           ...................-------------------..........T...............................
Z.bailii_9           ...................-------------------.......... ...............................
Z.bailii_10          ...................-------------------.......... ...............................
Z.bailii_12          ...................-------------------.......... ...............................
Z.bailii_13          ...................-------------------.......... ...............................
```

Figure 4.5: ClustalW alignment of the *Zygosaccharomyces bailii* strains rDNA region ITS2.

```
                           10        20        30        40        50        60        70        80        90       100
                   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii_LEU2_ISA_1307     ATGTCTAAAAACATCGTTGTTCTACCAGGTGACCACGTTGGTCAAGAAATCGCACAAGAAGCTATTAAAGTGCTGGAGGCGATCTCTGAAGTCAGCCCTA
Z.rouxii_LEU2_NCYC_568     .....A.............C............T.CC............T.......C..C.....T.A.....C...........T..T...G

                          110       120       130       140       150       160       170       180       190       200
                   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii_LEU2_ISA_1307     AGGCCAAATTCAACTTTCAGCATCATTTGATCGGGGTAGTGCAATTGATGCAACTGGATCTCCTTTGCCGGATGAAGCTCTTGCGGCTGCTAAGAAGGC
Z.rouxii_LEU2_NCYC_568     ..........T........A..C.....A..T..T...GC...T.......................A..A.....T........T..........A..

                          210       220       230       240       250       260       270       280       290       300
                   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii_LEU2_ISA_1307     CGACGCTGTGCTTCTCGGTGCCGTCGGCGGTCCTAAATGGGGTACAGGCGCCGTTAGGCCTGAGCAAGGTTTGCTAAAGATTCGTAAGGAATTGCAATTG
Z.rouxii_LEU2_NCYC_568     A..T..A.....AT.A.....T..T..T.....A...........C..TT.......A.....A......C..........C.................

                          310       320       330       340       350       360       370       380       390       400
                   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii_LEU2_ISA_1307     TACGCAAACTTGAGACCTTGCAACTTTGCTTCTGAATCTTTACTCGAACTTTCTCCTTTGAAACCACAACATGCCAGAGGTACAGATTTTATCGTTGTTA
Z.rouxii_LEU2_NCYC_568     .....C..T........A...........A.........C....T..T.....A..A.....G.....G.....A.AG..A..C......G.T.....A.

                          410       420       430       440       450       460       470       480       490       500
                   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii_LEU2_ISA_1307     GAGAACTAGTTGGAGGTATCTACTTCGGGCAGCGTAAAGAAGACGATGGTGACGGTGTCGCTTGGGACAGTGAAAGGTACACCAAGCCAGAAGTCCAGCG
Z.rouxii_LEU2_NCYC_568     .....T....G..C..............TG.A...........T..A.....T..C..T.................AA...........C.....T..A..

                          510       520       530       540       550       560       570       580       590       600
                   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii_LEU2_ISA_1307     TATCACTAGAATGGCTGCATTTTTAGCTCTTCAGCACAATCCTCCTCTGCCGATTTGGTCCCTTGACAAGGCCAATGTGCTGGCCTCCTCCAGATTATGG
Z.rouxii_LEU2_NCYC_568     .T.A..C........A.....C..G..C..A..A.....C..A.....A..T..C.....TT.G........T..C..CT....T..T......C.T...

                          610       620       630       640       650       660       670       680       690       700
                   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii_LEU2_ISA_1307     AGAAAGACAGTGGAGGAGACTATCAAGAACGAGTTCCCTCAATTGACAGTTAAACACCAATTGATTGACTCTGCCGCCATGATTTTGATTAAGAACCCAA
Z.rouxii_LEU2_NCYC_568     ........T..A..A..A..C..T.....T................A.C.C..C....................T...........AG.C....G.....

                          710       720       730       740       750       760       770       780       790       800
                   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii_LEU2_ISA_1307     CTCAGTTGAATGGTATCATTATTACTAGCAACATGTTTGAGGATATTATATCTGATGAAGCATCAGTAATTCCTGGTTCTCTCGGTCTGCTACCCTCGGC
Z.rouxii_LEU2_NCYC_568     ....A............G..T.A..A..T...T....C.GT..C..C..T........G..C..T..G..C..A.....G..T.....AT....A..A..

                          810       820       830       840       850       860       870       880       890       900
                   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii_LEU2_ISA_1307     TTCTTTGGCCTCTTTGCCAGACACGAATGAAGCCTTTGGTCTTTACGAGCCTTGTCATGGTTCGGCGCCTGATTTACCAAAGGGTAAAGTGAATCCTGTT
Z.rouxii_LEU2_NCYC_568     ...A.....T..............C.....G...........G.....A..C.....C.....T..T......C................T..C...A..

                          910       920       930       940       950       960       970       980       990      1000
                   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii_LEU2_ISA_1307     GCGATGATTCTTTCTGCAGCGATGATGTTGAAACTTTCTTTGAATATGGCCAAAGAGGGAGAGGCTCTCGAAACTGCCGTAAAGCAGGTTCTAGATTCGG
Z.rouxii_LEU2_NCYC_568     ..T......T.A.....C..T..................A.....CT.AT....G..A..T..A...G.G..G.AA..A..T......A...........C.

                          1010      1020      1030      1040      1050      1060      1070      1080      1090
                   ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|...
Z.bailii_LEU2_ISA_1307     GTGTTAGAACTGGCGACTTGGGCGGTTCCAATAGCACCTCCGAGGTTGGCGATGCCATCGCCAAGGCTGTAAAGCAAATTTTGGCTTGAACCA
Z.rouxii_LEU2_NCYC_568     .A..G........T..T..A..T.....T..C.....T..T..A..C..T..C..............A..C...G....C.....A...T
```

Figure 4.6: ClustalW alignment of *Z. bailii* and *Z. rouxii LEU2* ORF. NCBI accession numbers listed in table 2.5

```
                                   10        20        30        40        50        60        70        80        90       100
                          ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii_HIS3_ATCC_36947  ---------ATGATAGGAGAACACAAGGCCCTTGTCCAGAGGATTACCAACGAGACCAAAATTCAGATTGCCATCTCGCTAAATGGTGGCCATATAGAGC
Z.rouxii_HIS3_NCYC_588    ATGATTCCTGA.CA...T..G........TT....G.....A.A.........A........C.....C.....T..CT..........T..C..T..AA

                                  110       120       130       140       150       160       170       180       190       200
                          ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii_HIS3_ATCC_36947  TTCCAGAATCGATATTGGACAAAAATCCAACGCAGGAAAATAACATAGCTACGCAGGCCACCACTTCTCAAGTGATTGATATTCATACTGGTGTTGGATT
Z.rouxii_HIS3_NCYC_588    .......G..C..CA.A.GT..G..GAG.GTTG.AAGTG..GG.G..CG......A..T..T.G.......AC...C..C........C.....C.....

                                  210       220       230       240       250       260       270       280       290       300
                          ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii_HIS3_ATCC_36947  CCTGGATCATATGATTCATGCCCTAGCGAAGCACTCGGGTTGGTCTTTAATAGTCGAATGCATCGGCGATTTGCATATTGATGACCACCACACCACTGAG
Z.rouxii_HIS3_NCYC_588    ...C..C..............T..G.....A.....T........AC....C..T.....T..T..T........C..C.....T.....T........A

                                  310       320       330       340       350       360       370       380       390       400
                          ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii_HIS3_ATCC_36947  GATTGTGGCATTGCTCTCGGTGAAGCATTCAAACAAGCCCTTGGTCAAGTACGTGGTGTGAAAAGATTCGGTTGTGGATTTGCGCCCTTGGATGAAGCGT
Z.rouxii_HIS3_NCYC_588    ..C..C..T.....C..A.....T..T........G...T.G..A..G....................T..A.TC..T........A..A.....G..T.

                                  410       420       430       440       450       460       470       480       490       500
                          ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii_HIS3_ATCC_36947  TATCGCGTGCTGTCGTTGACTTGTCTAATAGACCTTATGCGGTGATTGACCTTGGTTTGAAAAGGGAGAAAATTGGAGACTTATCCTGTGAAATGATACC
Z.rouxii_HIS3_NCYC_588    ....AA.A.....T..C..TC.A..C........A...T.T..T.....AT.A..A........A..A.....C..T..T..G..............T..

                                  510       520       530       540       550       560       570       580       590       600
                          ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii_HIS3_ATCC_36947  TCATTTCCTCGAGAGTTTTATTGAAGCTGCAAGGCTGACGGTCCATATTGACTGCCTCAGAGGCTTTAACGACCACCACAGAAGTGAGAGTGCCTTTAAG
Z.rouxii_HIS3_NCYC_588    A.....T..A..A.......C......A..T..AT....T..T...G.G.....TT.G.....T..C........T...........A.....A......

                                  610       620       630       640       650       660       670       680
                          ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|.
Z.bailii_HIS3_ATCC_36947  GCACTGGCCGTAGCTATCAGAGAAGCTACTGCTCCCAATGGTACCAACGACGTCCCTTCAACAAAGGGCGTGCTAATGTAG
Z.rouxii_HIS3_NCYC_588    .....A..G.....CC.T.......A...T.A..T..............T.....A..T..G.....T..T..C.....GA
```

Figure 4.7: ClustalW alignment of *Z. bailii* and *Z. rouxii HIS3* ORF. NCBI Accession numbers listed in table 2.5

Genus or species	Primer name	Primer sequence, 5' to 3'	Tm (°C)	G.C (%)	Length (bp)
		Mid-*HIS3* sequencing			
Genus	ZhisF2	GAT YGA YAT TCA TAC YGG TGT YGG	53 - 54	38 - 54	24
Genus	ZhisR2	GCA CTY TCA CTT CTG TGR TGG TCG	58 - 59	50 - 58	24
		HIS3 5' end sequencing			
Z. bisporus	Specific reverse	TCG CCA ATG CAC TCG ACG ATC AAA G	62	52	25
	Degenerate forward F5	GMW AGG WGA RCA CAA GGC YYT TG	54 - 63	52	23
Z. kombuchaensis	Specific reverse	TCA CCA ATA CAC TCG ACG ATC AAT G	57	44	25
	Degenerate forward F5	GMW AGG WGA RCA CAA GGC YYT TG	54 - 63	52	23
Z. lentus	Specific reverse	TCG CCG ATA CAC TCC ACG ATC AAT G	61	52	25
	Degenerate forward F5	GMW AGG WGA RCA CAA GGC YYT TG	54 - 63	52	23
Z. mellis	Specific reverse	TCN CCN ATA CAC TCA ACG ATC AGT G	56 - 61	48	25
	Degenerate forward F6	CYY TTG TSC AGA GRA WTA CCA ACG	54 - 60	48	24
HIS3 3' end sequencing					
Z. bisporus	Specific forward	CAT TTC TTA GAG AGT TTC ACC GAG G	55	44	25
	Degenerate reverse R6	YYA CAT KAG MAC RCC CTT YGT WG	50 - 63	48	23
Z. kombuchaensis	Specific forward	CAT TTT CTN GAG AGT TTT ACC GAG G	53 - 56	42	25
	Degenerate reverse R6	YYA CAT KAG MAC RCC CTT YGT WG	50 - 63	48	23
Z. lentus	Specific forward	CAT TTT CTC GAG AGT TTT ACC GAG G	56	44	25
	Degenerate reverse R6	YYA CAT KAG MAC RCC CTT YGT WG	50 - 63	48	23
Z. mellis	Specific forward	CAT TTC CTT GAG AGT TTT ACT GAA G	52	36	25
	Degenerate reverse R6	YYA CAT KAG MAC RCC CTT YGT WG	50 - 63	48	23

Table 4.1: Degenerate and specific primers that amplified the mid-section, 5' end and then 3' end of the *Zygosaccharomyces HIS3* gene.

***Z. bailii* complete *HIS3* sequencing primers**		**Tm (°C)**	**G.C (%)**	**Length (bp)**
Z. bailii F1	AGA ACA CAA GGC CCT TGT CCA GAG G	63	56	25
Z. bailii R1	CCC TTT GTT GAA GGG ACA TCG TTG G	60	52	25

Table 4.2: Primer sequences for amplifying and sequencing the *Z. bailii* complete *HIS3* gene.

Genus / species	Amplicon	Size (bp)
Zygosaccharomyces	mid-HIS3	442
Z. bisporus	5'-HIS3	259
Z. kombuchaensis	5'-HIS3	259
Z. lentus	5'-HIS3	259
Z. mellis	5'-HIS3	235
Z. bisporus	3'-HIS3	181
Z. kombuchaensis	3'-HIS3	181
Z. lentus	3'-HIS3	181
Z. mellis	3'-HIS3	181

Table 4.3: Expected amplicon sizes for all *Zygosaccharomyces HIS3* sequencing primers.

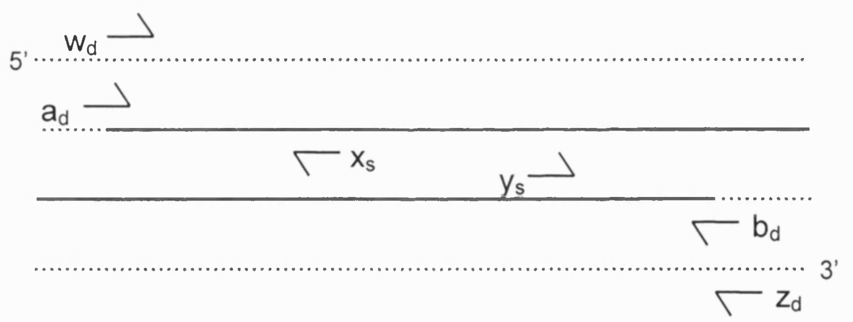

Figure 4.8: Diagram showing the methodology used to sequence the 5' and 3' ends of the *Zygosaccharomyces HIS3* gene. The central portion was amplified and sequenced with degenerate primers a_d and b_d. The 5' end was then amplified and sequenced with degenerate forward primer w_d and species-specific reverse primer x_s. The 3' end was amplified and sequenced with species-specific forward primer y_s and degenerate reverse primer z_d.

```
                            10        20        30        40        50        60        70        80        90       100
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.rouxii_NCYC_588   TTTGTGCAGAGAAATACCAACGAAACCAAAATCCAGATCGCCATTTCCTTAAATGGTGGTCACATTGAAATTCCAGAGTCCATCATAGGTAAGAAGAGAG
Z.mellis_3          .....C.....G...........G...........A........C.....G.....C........A.........A.T..G...........AG....G.
Z.bisporus_3        C....CA......TC..T.....G..............T.T...C..G..............T..A..GG..........A..TT.G.........CCCA
Z.lentus_1          .....CA....G.TA..G..T........G.....A..T..G..C..GC.......A..C.....AA.G...........G..AT.GA.CG.....CC..
Z.kombuchaensis_1   C....CA.A..G.TC..G..T..G..T..G.....A..T..G..C..AC.......A..C.....AA.G...........G..AT.GA........CC..
Z.bailii_ATCC_36947 C....C.....G.T.........G........T.....T.....C..GC..........C..T..A..GC.......A..G..AT.G.AC..A..TCC.A
                           110       120       130       140       150       160       170       180       190       200
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.rouxii_NCYC_588   TTGAAAGTGATGGCGTACGTACGCAAGCTACTAGTTCTCAAACGATCGACATTCATACCGGTGTCGGATTCCTCGACCATATGATTCATGCTCTGGCGAA
Z.mellis_3          ....G..C.....T...GCC..A............G....GTT...T.....C..C..A........T................................
Z.bisporus_3        CGC..GAAA..A..A..GC...A..G..C..CTC......GT....T..TG.C.....A.....T......T.G..............C..A..A.....
Z.lentus_1          ACC.GGAA..C......GC...A..G..C...G.C.....GT....T..T..C..C..T..C........TT.G..............C..G..A.....
Z.kombuchaensis_1   ACC.GGAA......A..GC......G..C...G.C..C..GGTC..T..T..C..C..T..C.....G..TT.G...........C..C..G..C.....
Z.bailii_ATCC_36947 CGC.GGAAA..AA.A..GC......G..C..C.C......GT....T..T........T.....T........G..T..............C..A.....
                           210       220       230       240       250       260       270       280       290       300
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.rouxii_NCYC_588   ACACTCTGGTTGGTCACTAATCGTTGAATGTATTGGTGATTTGCACATCGATGATCACCATACCACTGAAGACTGCGGTATTGCCCTAGGTGATGCTTTC
Z.mellis_3          ......A...........G........G................................C.......................................
Z.bisporus_3        ......C..G.....TT.G.....C..G..C.....C..............C..C.....C.....A..G.....T........G........A..A...
Z.lentus_1          G.....C.........T.G.....G..G.....C..C...........T...........C..............T...........T.....A......
Z.kombuchaensis_1   G.....C.........T.G.....C..G.................T..............C..T...........T...........T.....A......
Z.bailii_ATCC_36947 G.....G........TT....A..C.....C..C..C........T..T.....C.....C........G..T..T...C....T..C.....A..A...
                           310       320       330       340       350       360       370       380       390       400
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.rouxii_NCYC_588   AAACAGGCCTTGGGACAGGTACGTGGTGTGAAAAGATTTGGATTCGGTTTTGCGCCATTAGATGAGGCTTTATCAAGAGCTGTTGTCGATCTATCCAATA
Z.mellis_3          .G...A...C..........G..............G.....C.....G.....T..T........A..G..............A......T.G..T..C.
Z.bisporus_3        ..G......C.T..C..A..C..............G.....T.G...A.....A..TC.G.....A........GC.T.....C..G.....G.....C.
Z.lentus_1          ........TC.A..C..A..G....................T..T..G........TC.G..C......C....CC.T..A..C......T.G.....C.
Z.kombuchaensis_1   ........TC.T..C..A..G..C..............C..T..T..G.....A..CC.T..............GC.T..C..C......T.G.....C.
Z.bailii_ATCC_36947 .....A...C.T..T..A....................C..T.GT..A........C..G.....A..G.....GC.T.....C..T..CT.G..T....
                           410       420       430       440       450       460       470       480       490       500
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.rouxii_NCYC_588   GACCATATTCTGTTATTGAATTAGGATTGAAAAGAGAAAAAATCGGTGATTTGTCCTGTGAAATGATTCCACATTTTCTAGAAAGTTTTACTGAAGCAGC
Z.mellis_3          ......G...............G.....A.............................C........C........C..T..G..............G..
Z.bisporus_3        ....C...CG......A..TC....C........G.....G..T.....C.....T..C........C..C.....CT....G.....C..C..G..T..
Z.lentus_1          .G.....CG.A........C.....T........G..G...........CC..........G...G.C..T........C..G........C..G..T..
Z.kombuchaensis_1   .G..T..CG.A..G..CA.C..G..C...........G...........CC..........G.....C..T........C..G........C..G..T..
Z.bailii_ATCC_36947 ....T...G.G..G.....CC.T..T........G..G.....T..A..C..A..............A..T.....C..C..G.......T......T..
                           510
                    ....|....|....|.
Z.rouxii_NCYC_588   TAGATTGACTGTTCAT
Z.mellis_3          C........C..C...
Z.bisporus_3        C..G.....AA.....
Z.lentus_1          G...C....CA.C...
Z.kombuchaensis_1   G...C....AA.....
Z.bailii_ATCC_36947 A..GC....G..C...
```

Figure 4.9: ClustalW alignment of the *Zygosaccharomyces species* partial *HIS3* gene.

```
                    10        20        30        40        50        60        70        80        90       100
            ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii_T  GGATTCCTGGATCATATGATTCATGCCCTAGCGAAGCATTCGGGTTGGTCNTTAATAGTTGAATGCATTGGTGATTTGCATATTGATGACC      CCA
Z.bailii_1  ......................................C....................C........C..C...................ACCACA...
Z.bailii_2  .............................T..A..........................................................ACCACA...
Z.bailii_3  ......................................C....................C........C..C...................ACCACA...
Z.bailii_4  ......................................C....................C........C..C...................ACCACA...
Z.bailii_5  ......................................C....................C........C..C...................ACCACA...
Z.bailii_8  ......................................C....................C........C..C...................ACCACA...
Z.bailii_9  ...........................................................................................ACCACA...
Z.bailii_13 ...........................................................................................ACCACA...
                   110       120       130       140       150       160       170       180       190       200
            ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii_T  CTGAGGATTGTGGCATTGCTCTCGGCGAAGCATTCAAACAAGCCCTTGGCCAAGTCCGTGGCGTGAAAAGATTCGGTTGTGGATTTGCGCCCTTGGATGA
Z.bailii_1  .................................................T...........T......................................
Z.bailii_2  ...........................................T.....T......................................A..T........
Z.bailii_3  .............................................................T......................................
Z.bailii_4  .............................................................T......................................
Z.bailii_5  .................................................T.....A.....T......................................
Z.bailii_8  .................................................T...........T......................................
Z.bailii_9  ....................................................................................................
Z.bailii_13 ....................................................................................................
                   210       220       230       240       250       260       270       280       290       300
            ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii_T  AGCGTTATCGCGTGCTGTCGTTGACTTGTCCAATAGACCTTATGCAGTGATTGACCTTGGTTTGAAAAGGGAGAAAATTGGAGACTTATCCTGCGAAATG
Z.bailii_1  .............................................G......................................................
Z.bailii_2  ........................................................................A...........................
Z.bailii_3  .............................................G......................................................
Z.bailii_4  ..............................T.....................................................................
Z.bailii_5  ....................................................................................................
Z.bailii_8  .............................................G......................................................
Z.bailii_9  ....................................................................................................
Z.bailii_13 ....................................................................................................
                   310       320       330       340       350       360       370       380       390
            ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|.
Z.bailii_T  ATACCTCATTTTCTCGANAGTTTCATTGAAGCTGCCAGGCTGANGGTCCATATTGACTGCCTCAGAGGCTTTAACGACCACCACAGAAGTGAAAGT
Z.bailii_1  ...........C....................................................................................
Z.bailii_2  .......................T..............A...................................T.....................
Z.bailii_3  ...........C................................................................................G...
Z.bailii_4  ...........C................................................................................G...
Z.bailii_5  ...........C...........T....................................................................G...
Z.bailii_8  ...........C................................................................................G...
Z.bailii_9  ..........................................................................T.....................
Z.bailii_13 ..........................................................................T.....................
```

Figure 4.10: ClustalW alignment of the *Zygosaccharomyces bailii* strains partial *HIS3* sequence.

4.2.3 Specific primer design

Using the *HIS3* ClustalW alignment as a reference, many forward and reverse primers were designed to be tested for specificity to each *Zygosaccharomyces* species (see tables 4.4 and 4.5 for sequences and amplicon sizes).

Zygosaccharomyces	Primer sequence, 5' to 3'	Tm (°C)	G.C (%)	Length (bp)
Z. bailii				
Zba F1	GCC CTA GCG AAG CAC TCG GGT TGG	64	66	24
Zba F2	TAT TGA TGA CCA CCA CAC CAC TGA G	57	48	25
Zba F3	CTG AGG ATT GTG GCA TTG CTC TCG	61	54	24
Zba F4	TCA TAT GAT TCA TGC CCT AGC GAA G	58	44	25
Zba R1	TAG ACA AGT CAA CGA CAG CAC GCG	61	54	24
Zba R2	CTT TTC AAA CCA AGG TCA ATC ACC G	56	44	25
Zba R3	GAG GAA ATG AGG TAT CAT TTC ACA G	54	40	25
Zba R4	CAC CGC ATA AGG TCT ATT AGA CAA G	56	44	25
Zba R5	GCA CGC GAT AAC GCT TCA TCC AAG	61	54	24
Zba R6	TGC AGC TTC AAT AAA ACT CTC GAG	57	41	24
Zba R7	ATC CAA GGG CGC AAA TCC ACA ACC G	62	56	25
Z. bisporus				
Zbi F1	CAG AGT CAA TTT TGG GTA AGA AG	51	39	23
Zbi F2	GTA AGA AGC CCA CGC AAG AAA TAG	56	46	24
Zbi F3	GAG GTT CCA GAG TCA ATT TTG GG	54	48	23
Zbi R1	AGC CTA GAT CTA TAA CAG CGT AG	55	43	23
Zbi R2	GGT GAA ACT CTC TAA GAA ATG GGG G	56	48	24
Zbi R3	TAT AAC AGC GTA GGG TCT GTT GGA CAG	60	48	27
Zbi R4	AAT TGT CAA CCT GGC AGC CTC GGT G	62	56	25
Zbi R5	AGC TTC ATC CAG AGG TGC AAA TCC G	61	52	25
Zbi R6	TAG ATC TAT AAC AGC GTA GGG TCT GTT G	59	43	28
Z. kombuchaensis				
Zko F1	CGA GTG TAT TGG TGA TTT GCA TAT CG	57	42	26
Zko F2	ATA GCT ACG CAG GCC ACT GGC TCC CAG	67	63	27
Zko F3	CAA AAG GAT CAC GAA TGA GAC TAA G	55	40	25
Zko R1	GCC CAA GTT GAT CAC TGC GTA AG	57	52	23
Zko R2	CGA TAA AGC CTC ATC AAG GGG TG	57	52	23
Zko R3	GGC ACG CGA TAA AGC CTC ATC AAG G	62	56	25
Z. lentus				
Zle F1	GGA GTG TAT CGG CGA TTT GCA CAT TG	60	50	26
Zle F2	CAC AGG CCA CTG GCT CTC AAG TG	60	61	23
Zle F3	CAA GAG GAT AAC GAA TGA AAC CAA G	55	40	25
Zle R1	ACC TAA GTC AAT AAC TGC GTA TG	52	49	23
Zle R2	GGA TAG AGC CTC GTC CAG AGG CG	62	65	23
Zle R3	TGC ACG GGA TAG AGC CTC GTC CAG AG	65	62	26

Table 4.4: Sequences of all potentially species-specific *Zygosaccharomyces* primers.

Zygosaccharomyces	Primer sequence, 5' to 3'	Tm (°C)	G.C (%)	Length (bp)
Z. mellis				
Zme F1	GCC ACA CAA GCT ACT AGT GCT CAA GTT	61	48	27
Zme F2	GGT TGA GAG CGA TGG TGT AGC CAC A	61	56	25
Zme R1	CCT TTT CAC ACC ACT GCA CCG	56	57	21
Zme R2	TGC ACC GTC CCA GGG CTT GTC	59	67	21
Z. rouxii				
Zro F1	AGT TGA AAG TGA TGG CGT ACG	54	48	21
Zro F2	GCA AGC TAC TAG TTC TCA AAC G	54	48	22
Zro F3	AGG TAA GAA GAG AGT TGA AAG T	50	36	22
Zro R1	AGA ATA TGG TCT ATT GGA TAG	45	33	21
Zro R2	TGC TTC AGT AAA ACT TTC TAG A	50	32	22

Table 4.4 (*continued*): Sequences of all potentially species-specific *Zygosaccharomyces* primers.

		Expected amplicon sizes for each *Zygosaccharomyces* specific primer pair tested						
Species	Primer	R1	R2	R3	R4	R5	R6	R7
Z. bailii	F1	207	245	300	225	191	321	176
	F2	151	184	235	164	135	256	118
	F3	132	165	211	145	116	232	99
	F4	220	258	313	238	204	334	189
Z. bisporus	F1	354	419	344	440	296	350	
	F2	338	404	329	425	341	335	
	F3	361	426	351	447	303	357	
Z. kombuchaensis	F1	202	151	157				
	F2	312	261	267				
	F3	418	367	373				
Z. lentus	F1	202	151	157				
	F2	314	263	296				
	F3	418	367	373				
Z. mellis	F1	210	196					
	F2	229	215					
Z. rouxii	F1	313	400					
	F2	289	384					
	F3	325	412					

Table 4.5: Expected amplicon sizes for all *Zygosaccharomyces* potentially specific primer pairs.

4.2.4 Specific primer testing

Initial screening

In a similar method to *Brettanomyces* / *Dekkera* primers, all combinations of primer were then tested on pairs of *Zygosaccharomyces* species to rapidly screen which were immediately non-specific and which could potentially be specific for each species. PCR conditions were the same generic conditions described in section 2.3.1. The methodology has been described in section 3.4.3 and so to avoid repetition, not all results are included in this section. Typical results included: 1. The primer pair only amplified the target species and no others (figure 4.11A) - these primers were deemed provisionally specific, 2. The primer pair amplified the target species and some or all of the other species (figure 4.11B), 3. The primer pair amplified unexpected products (figure 4.11C), 4. The primer pair amplified nothing (figure 4.11D) and 5. Combinations of 1 to 4.

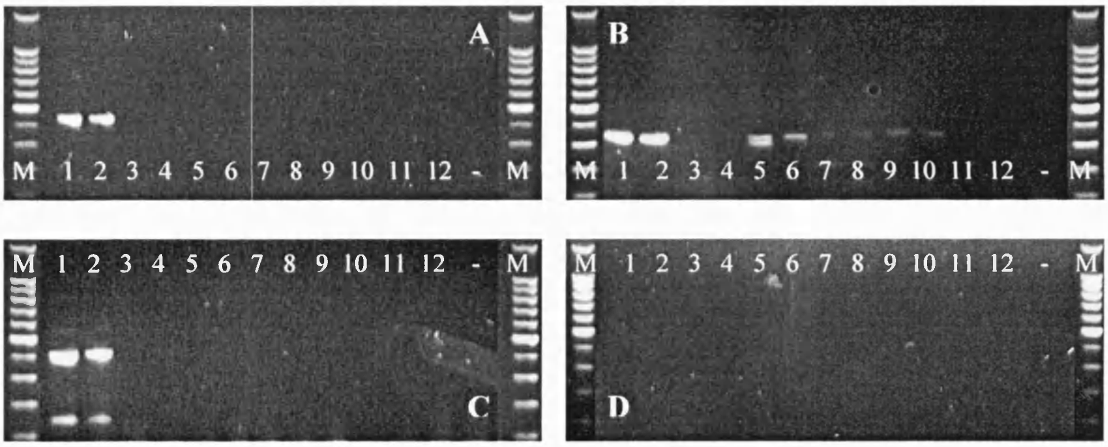

Figure 4.11: Potentially *Z. bisporus* specific primers (A) F3 / R2 (B) F2 / R3 (C) F2 / R4 (D) F1 / F1 with (1) *Z. bisporus* 1, (2) *Z. bisporus* 3, (3) *Z. bailii* 1, (4) *Z. bailii* 2, (5) *Z. rouxii* 1, (6) *Z. rouxii* 2, (7) *Z. mellis* 1, (8) *Z. mellis* 3, (9) *Z. lentus* 1, (10) *Z. lentus* 2, (11) *Z. kombuchaensis* 1 and (12) *Z. kombuchaensis* 2.

Target species

After the initial screen of each primer pair against two strains of each *Zygosaccharomyces* species, the pairs that were provisionally specific (as demonstrated with *Z. bisporus* F3 / R2 in figure 4.11A) were then tested on all available target species. Typical results included: 1. All target isolates were successfully amplified (figure 4.12A) and 2. Amplification was not consistent amongst isolates of the same species (figure 4.12B).

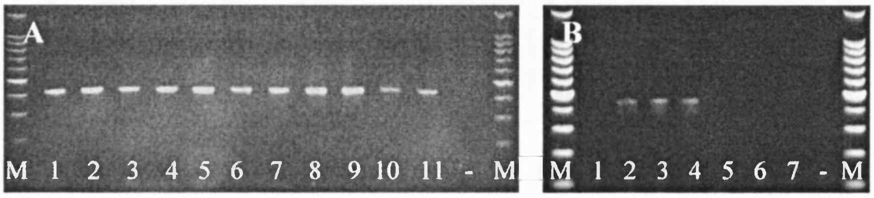

Figure 4.12: Examples of typical results when provisionally specific primer pairs are tested on all target isolates. (A) *Z. rouxii* F3 / R2 (1 – 11) *Z. rouxii* 1 –11. (B) *Z. bisporus* F1 / R4 (1 – 7) *Z. bisporus* 1 – 7.

Other species

Primer pairs that had passed the preliminary test and then shown to amplify all target isolates were tested on all available isolates of the closest sibling. Typical results included: 1. No sibling species were amplified (figure 4.13A), 2. The primer pair amplified all or some of the sibling species isolates (figure 4.13B), and 3. Due to the close relationship between siblings, some isolates had been misidentified by physiological methods and the primers highlighted this (figure 4.13C). Definitive identification was then performed via D1 / D2 sequencing.

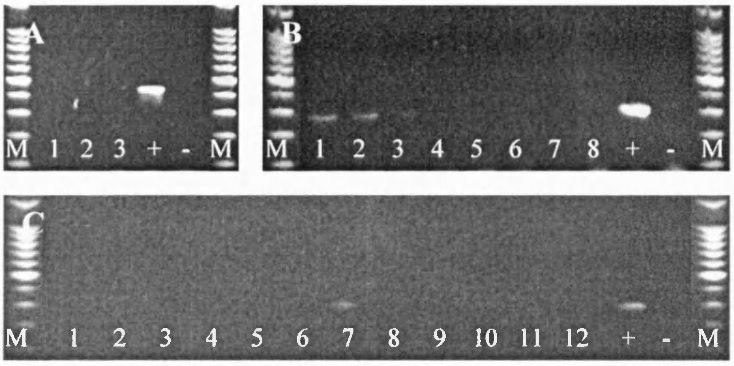

Figure 4.13: Examples of typical results when provisionally specific primer pairs are tested on their targets closest sibling. (A) *Z. rouxii* F3 / R2 (1 – 3) *Z. mellis* 1 – 3. (B) *Z. kombuchaensis* F2 / R1 (1 – 8) *Z. lentus* 1 – 8. (C) *Z. bisporus* F3 / R4 (1 – 6) *Z. bailii* 1 – 6 (7) *Z. bisporus* 8 (strain previously classified as *Z. bailii*, D1 / D2 sequencing after this result showed it was actually *Z. bisporus*) (8 – 12) *Z. bailii* 7 – 11.

Primer pairs that were still potentially specific at this stage were then tested on the rest of the *Zygosaccharomyces* species. Most primer pairs did not amplify any of the other species, as they were too dissimilar in sequence, especially as these primers had not amplified the closest sibling. However, some primer pairs did amplify more distantly related species rather than closest siblings (figure 4.14).

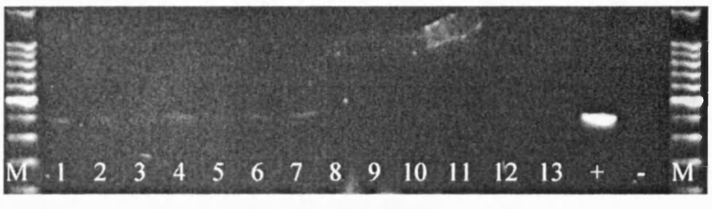

Figure 4.14: Example of a provisionally specific primer pair amplifying unexpected species. *Z. bisporus* primer pair F3 / R2 (1 – 10) *Z. rouxii* 1 –10, (11 – 13) *Z. mellis* 1 – 3.

Outgroups

Primer pairs that had been species-specific to this point were then tested on a selection of outgroup species. For *Zygosaccharomyces*, the outgroup species were chosen as those in the genera *Zygotorulaspora*, *Torulaspora* and *Lachancea* (clades 8, 9 and 10 from [133] and are listed in table 4.6. Each specific primer pair was tested on two representative species of each genus.

Clade	Species
8	*Zygotorulaspora florentinus* (Formerly *Zygosaccharomyces*)
8	*Zygotorulaspora mrakii* (Formerly *Zygosaccharomyces*)
9	*Torulaspora globosa*
9	*Torulaspora franciscae*
9	*Torulaspora pretoriensis*
9	*Torulaspora delbrueckii*
9	*Torulaspora microellipsoides* (Formerly *Zygosaccharomyces*)
10	*Lachancea cidri* (Formerly *Zygosaccharomyces*)
10	*Lachancea fermentati* (Formerly *Zygosaccharomyces*)
10	*Lachancea thermotolerans* (Formerly *Kluyveromyces*)
10	*Lachancea waltii* (Formerly *Kluyveromyces*)

Table 4.6: List of species in clades 8, 9 and 10 [133]

Z. bailii species-specific primer pairs F3 / R2 and F3 / R1 amplified some of the outgroup species (figure 4.15A). F3 / R1 and F3 / R2 gave identical results and produced oversized amplicons with *Zygotorulaspora florentinus* and correctly sized false positive amplicons with *Lachancea cidri*. In contrast, these same species were negative when tested with *Z. bailii* primer pair F1 / R2 that had previously amplified members of the *Zygosaccharomyces* clade (figure 4.15B).

The *L. cidri* test strain D1 / D2 region was sequenced and was identical to that from the GenBank database. A partial *HIS3* gene sequence was then produced and aligned with the *Zygosaccharomyces* species (figure 4.16).

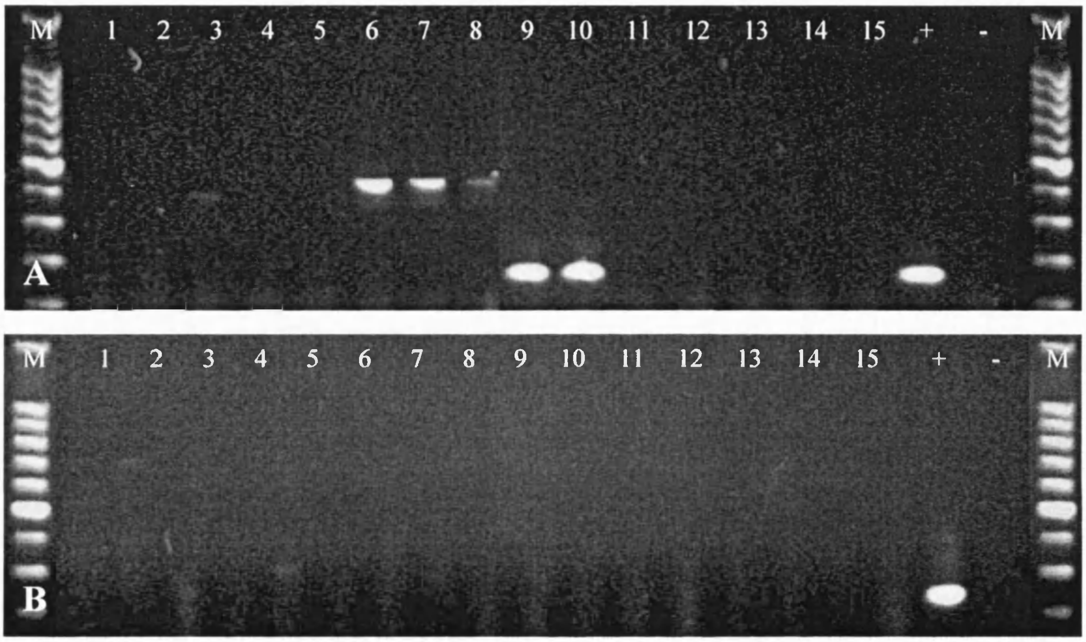

Figure 4.15: (A) *Z. bailii* specific primers F3 / R2 (B) *Z. bailii* non-specific primers F1 / R2. (1) *L. fermentati* 1, (2) *L. fermentati* 2, (3) *Zt. mrakii* 1, (4) *Zt. mrakii* 2, (5) *Zt. mrakii* 3, (6) *Zt. florentinus* 1, (7) *Zt. florentinus* 2, (8) *Zt. florentinus* 3, (9) *L. cidri* 1, (10) *L. cidri* 2, (11) *T. microellipsoides* 1, (12) *T. microellipsoides* 2, (13) *T. microellipsoides* 3, (14) *T. delbrueckii* 1, (15) *T. delbrueckii* 2.

```
                                     10        20        30        40        50        60        70        80        90       100
                            ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii HIS3               GCGAAGCACTCGGGTTGGTCTTTAATAGTCGAATGCATCGGCGATTTGCATATTGATGACCACCACACCACTGAGGATTGTGGCATTGCTCTCGGTGAAG
Z.bisporus HIS3             .....A.....C..G........G..C.....G.....T...........C..C..C..............A.....C.....T.....G..A.......
Z.lentus HIS3               ...........C........A..G..C..G..G..T..............C........T..............A..C.....T.....C..T.......
Z.kombuchaensis HIS3        ...........C........A..G..C.....G..T..T..T...........C.....T........T.....A..C.....T.....C..T.......
Z.rouxii HIS3               .....A.....T........AC....C..T.....T..T..T........C..C.....T.....T........A..C..C..T.....C..A.....T.
Z.mellis HIS3               .....A.....A........AC.G..C..T..G..T..T..T........C..C.....T..............A..C..C..T.....C..A.....T.
L.cidri_NCYC_1567           ..T.....T..A..C.....CC.C..T..A.....T..T..T..C........C..C..T..T.....T..C...........T...........A....

                                    110       120       130       140       150       160       170       180       190       200
                            ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii HIS3               CATTCAAACAAGCCCTTGGTCAAGTACGTGGTGTGAAAAGATTCGGTTGTGGATTTGCGCCCTTGGATGAAGCGTTATCGCGTGCTGTCGTTGACTTGTC
Z.bisporus HIS3             .......G..G........C.....C..............G..T.....C........A..TC..........T.................G..TC....
Z.lentus HIS3               .T........G..T..A..C.....G.................T....T...G........TC....C..G..TC....C.....A.....C..T.....
Z.kombuchaensis HIS3        .T........G..T.....C.....G..C...................T...G.....A...C.T.....G..T...........C.....C..T.....
Z.rouxii HIS3               .T........G...T.G..A..G....................T..A.TC..T........A..A.....G..T.....AA.A.....T..C..TC.A..
Z.mellis HIS3               .T....G.........G..A..G..G..............G..T..C.TC..G.....T..T..A..............AA.A.....A..C..T.....
L.cidri_NCYC_1567           .......GG....T..A...GCT..G........C........T..A..............AC.T..C..G..TC.C........G..T..G..T..A..

                                    210       220       230       240       250       260       270       280       290       300
                            ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii HIS3               TAATAGACCTTATGCGGTGATTGACCTTGGTTTGAAAAGGGAGAAAATTGGAGACTTATCCTGTGAAATGATACCTCATTTCCTCGAGAGTTTTATTGAA
Z.bisporus HIS3             C..C.....C..C..T..T..A..T..A..C...........A..G.....T.....G..T..C........C..C......T.A........C.CC..G
Z.lentus HIS3               C..C..G..A..C..A..T......T.A....................C..T...C.G........G...G.C........T.............CC..G
Z.kombuchaensis HIS3        C..C..G.....C..A.....CA..T.G..C........A........C..T...C.G........G.....C........T.............CC..G
Z.rouxii HIS3               C........A...T.T..T.....AT.A..A........A..A.....C..T..T..G..............T..A.....T..A..A.......C....
Z.mellis HIS3               ...C.....A.G.T.T..T.....AT.G..A..A.....A..A.....C..T..T..G.....C........C..A........T..........C....
L.cidri_NCYC_1567           C..C.....A.TC..A...........G..A.....G...........C..T..T.....GACA........T..A.....T..A..A..C..CGC...G

                                    310       320
                            ....|....|....|....|.
Z.bailii HIS3               GCTGCAAGGCTGACGGTCCAT
Z.bisporus HIS3             .....C...T....AA.T...
Z.lentus HIS3               .....G..A.....CA.....
Z.kombuchaensis HIS3        .....G..A.....AA.T...
Z.rouxii HIS3               ..A..T..AT....T..T...
Z.mellis HIS3               ..G..C..AT....C......
L.cidri_NCYC_1567           ..A..T..AA.T..TC.....
```

Figure 4.16: ClustalW alignment of the *Zygosaccharomyces species* and *L. cidri* partial *HIS3* gene.

4.2.5 Specific primer optimisation

In some cases, no primer pairs were specific after testing the complete *Zygosaccharomyces* genus and the outgroup species. As the primers had already been designed to encompass as many mismatches as possible and remain within the design parameters, the PCR conditions were altered in an attempt to improve specificity, as opposed to new primer being designed. *Z. lentus* primers provide such an example. The pair chosen for optimisation did not amplify *Z. kombuchaensis* (the closest sibling), as it was deemed that more divergent species would be simpler to eliminate by altering reaction conditions. The *Z. lentus* primer pair F2 / R1 had previously amplified most *Z. bisporus* isolates and a single *Z. bailii* isolate, albeit very weakly. A temperature gradient experiment was performed on a selection of target isolates and a *Z. bisporus* isolate to find an annealing temperature where all target isolates would be faithfully amplified and non-target isolates would not be picked up (figure 4.17). It can be seen that the target isolates are faithfully amplified up to a temperature of 63.5°C whilst the non-target isolate is only weakly amplified throughout and is lost completely by 60.5°C. A reaction at a new annealing temperature of 65°C was performed on all target isolates and all non-target isolates that had previously given a product (figure 4.18). The new annealing temperature amplifies all target species and does not amplify any of the non-targets. The *Z. lentus* F2 / R1 primers are therefore species-specific under these reaction conditions.

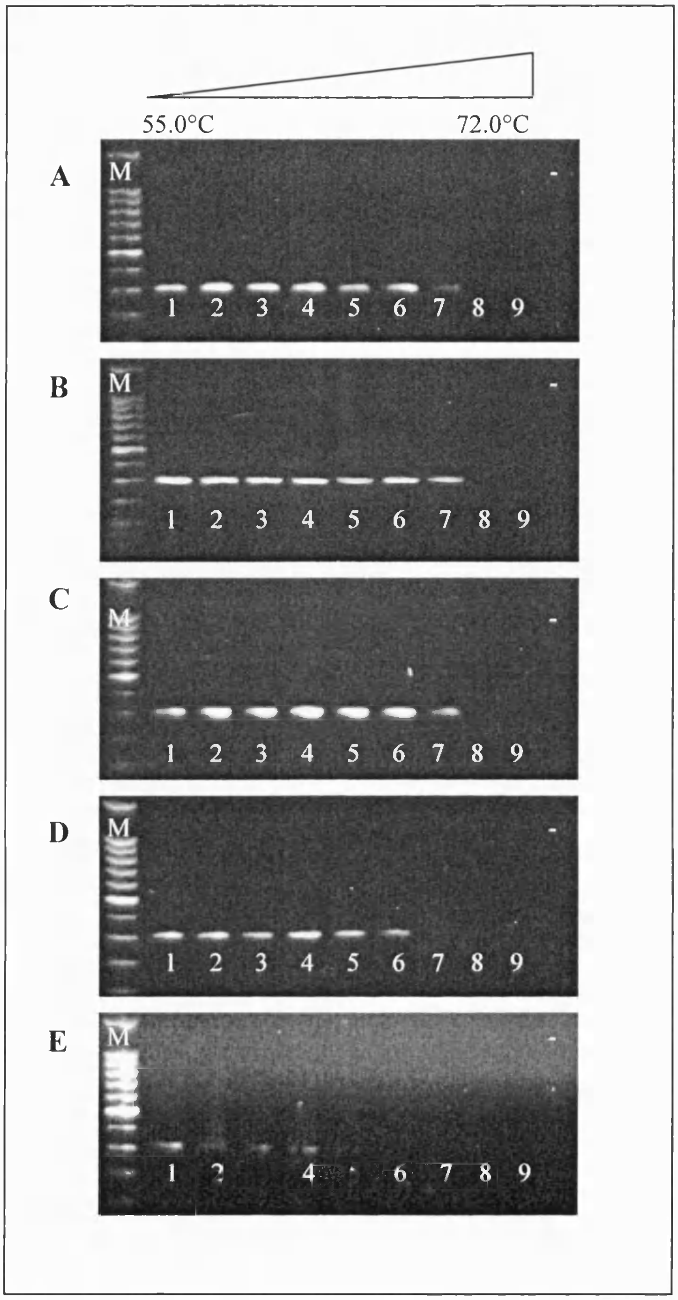

Figure 4.17: Annealing temperature gradient PCR with *Z. lentus* primers F2 / R1 and (A) *Z. lentus* 1, (B) *Z. lentus* 2, (C) *Z. lentus* 3, (D) *Z. lentus* 4 and (E) *Z. bisporus* 3. (1) 55.0°C (2) 55.2°C (3) 56.5°C (4) 58.2°C (5) 60.5°C (6) 63.5°C (7) 66.8°C (8) 69.6°C (9) 72.0°C

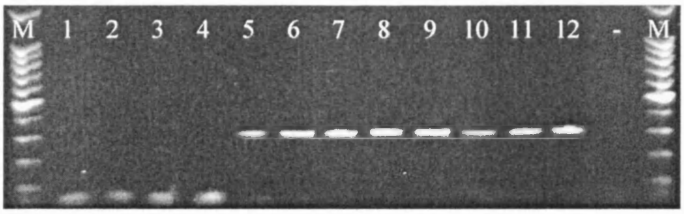

Figure 4.18: *Z. lentus* primers F2 / R1 at an annealing temperature of 65°C with target isolates and non-target isolates that are positive at 55°C. (1) *Z. bisporus* 1, (2) *Z. bisporus* 3, (3) *Z. bisporus* 4, (4) *Z. bailii* 7 and (5 – 12) *Z. lentus* 1 - 7

Other species that needed reaction optimisation were *Z. bailii* and *Z. kombuchaensis*. The two primer pairs that were *Z. bailii* specific when tested against *Zygosaccharomyces* species both also amplified *L. cidri*. As this was a very strong amplification, it was unlikely that this species would have been simple to eliminate using a temperature gradient. Instead, the *Z. bailii* F1 / R2 primer pair underwent temperature gradient optimisation (figure 4.19). The false *Z. rouxii* positive is eliminated at 56.8°C but the false *Z. lentus* positive persists, albeit weakly, up to 70°C.

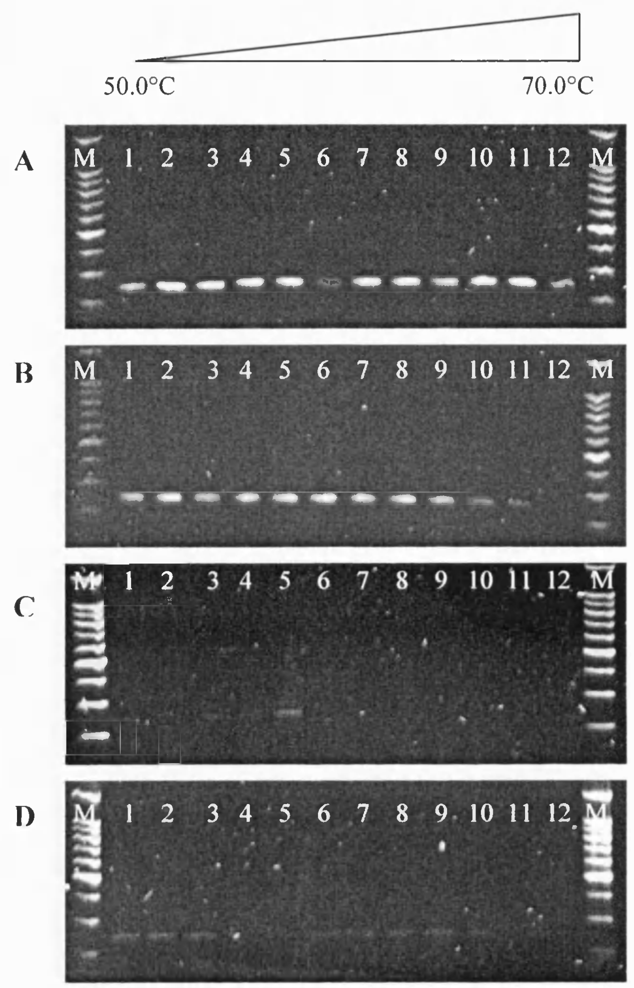

Figure 4.19: Annealing temperature gradient PCR with *Z. bailii* primers F1 / R2 and (A) *Z. bailii* 1, (B) *Z. bailii* 2, (C) *Z. rouxii* 3 and (D) *Z. lentus* 7 (1) 50.0°C (2) 50.5°C (3) 51.5°C (4) 53.2°C (5) 55.5°C (6) 58.4°C (7) 61.8°C (8) 64.6°C (9) 66.8°C (10) 68.4°C (11) 69.9°C (12) 70°C.

A Mg^{2+} ion concentration gradient PCR was then performed at 55°C annealing temperature to assess the effect on specificity (figure 4.20). The gradient was from 2.5mM to 1.0mM Mg^{2+} ion concentration, at 0.5mM intervals. The *Z. rouxii* isolate was not amplified at all, even at 2.5mM (the same conditions as the original tests). This isolate was only ever amplified very weakly and was consistently inconsistent showing the fidelity of primer binding was not strong. The *Z. lentus* false positive is eliminated between 2.5mM and 2.0mM Mg^{2+} ion concentration and the target isolate amplification becomes unreliable between 2.0mM and 1.5mM Mg^{2+} ion concentration.

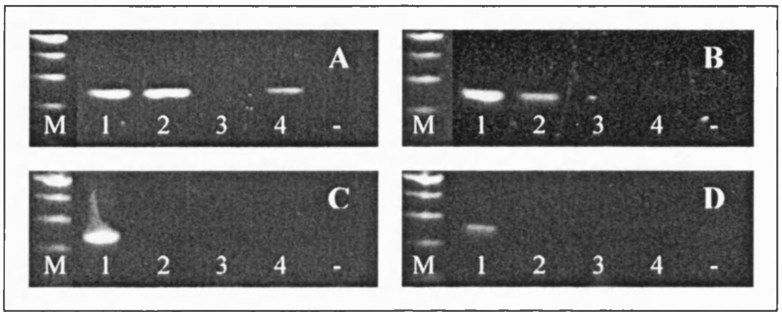

Figure 4.20: *Z. bailii* primers F1 / R2 with target isolates and non-target isolates that were positive at 55°C, at different Mg^{2+} ion concentration. (A) 2.5mM, (B) 2.0mM, (C) 1.5mM and (D) 1.0mM. (1) *Z. bailii* 1, (2) *Z. bailii* 2, (3) *Z. rouxii* 3 and (4) *Z. lentus* 7.

In order to investigate further the limits of *Z. bailii* primer binding to target isolates, another gradient was performed between 2.0mM and 1.5mM Mg^{2+} ion concentration, at intervals of 0.1mM (figure 4.21). It was estimated that a Mg^{2+} ion concentration of 1.8mM would faithfully amplify all targets while eliminating all non-targets. Furthermore, the annealing temperature would be increased by 5°C to 60°C. A reaction on all target isolates was performed under these conditions (figure 4.22) and they were faithfully amplified.

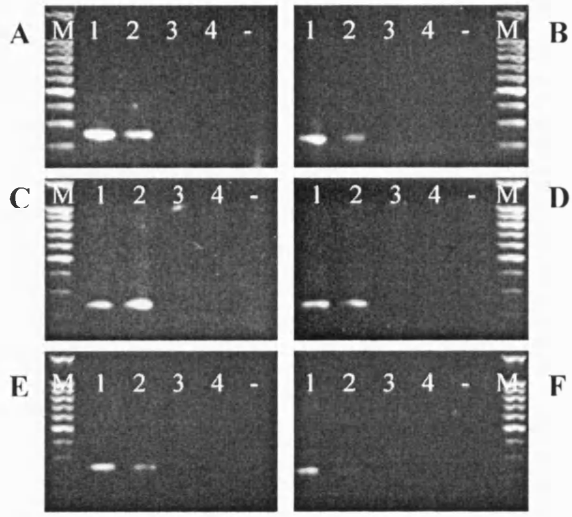

Figure 4.21: *Z. bailii* primers F1 / R2 with target isolates and non-target isolates that were positive at 55°C, at different Mg^{2+} ion concentration. (A) 2.0mM, (B) 1.9mM, (C) 1.8mM and (D) 1.7mM (E) 1.6mM (F) 1.5mM. (1) *Z. bailii* 1, (2) *Z. bailii* 2, (3) *Z. rouxii* 3 and (4) *Z. lentus* 7

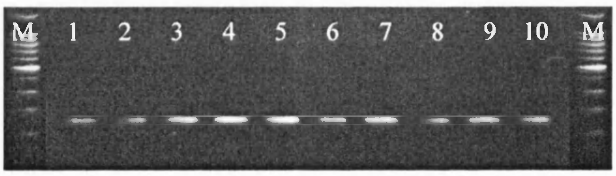

Figure 4.22: *Z. bailii* primers F1 / R2 with *Z. bailii* isolates under PCR conditions of 60°C annealing temperature and 1.8mM Mg^{2+} ions. (1 – 10) *Z. bailii* isolates 1 – 10.

All *Z. kombuchaensis* primer combinations amplified the closest sibling, *Z. lentus*. The *Z. kombuchaensis* primer pair F3 / R1 was optimised by annealing temperature gradient to be species-specific at 67.5°C (figure 4.23).

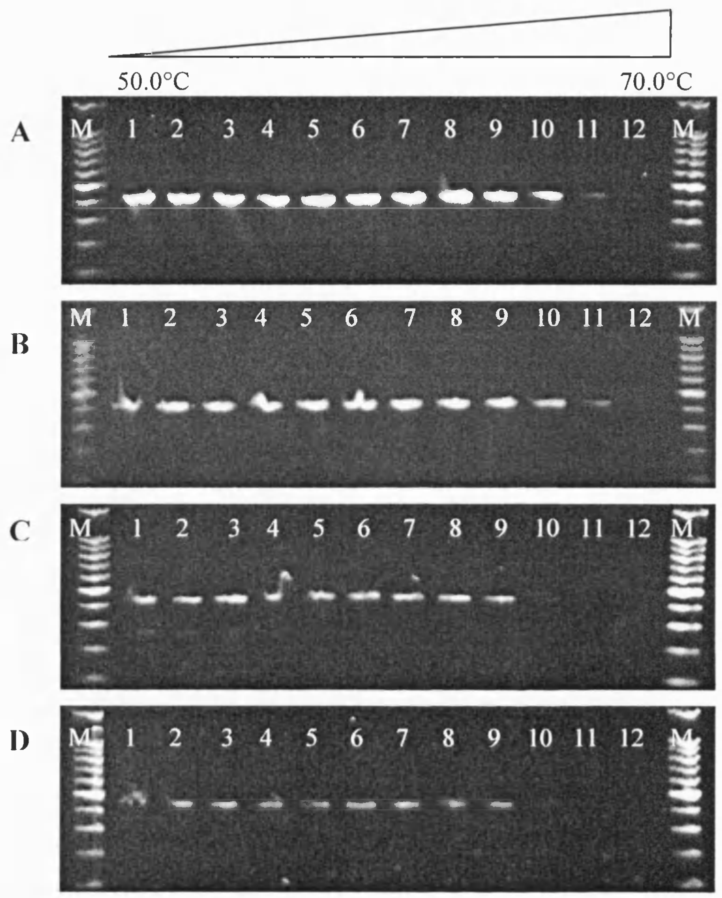

Figure 4.23: Annealing temperature gradient PCR with *Z. kombuchaensis* primers F3 / R1 and (A) *Z. kombuchaensis* 1, (B) *Z. kombuchaensis* 2, (C) *Z. lentus* 1 and (D) *Z. lentus* 2 (1) 50.0°C (2) 50.5°C (3) 51.5°C (4) 53.2°C (5) 55.5°C (6) 58.4°C (7) 61.8°C (8) 64.6°C (9) 66.8°C (10) 68.4°C (11) 69.9°C (12) 70°C.

4.3 Discussion

4.3.1 rDNA sequences

The *Zygosaccharomyces* D1 / D2 sequence alignment (figure 4.1) exhibits considerably less variation than that of the *Brettanomyces / Dekkera* species (figure 3.1). The variation is clustered in the same regions of the sequence, around the 100bp and 500bp nucleotide positions.

The most closely related pair of *Brettanomyces / Dekkera* species are *D. bruxellensis* and *D. anomala*; they have 38 single nucleotide polymorphisms (SNPs) between them within the approximately 591bp sequence alignment shown (this number does not include indels). The most closely related pair of *Zygosaccharomyces* species are *Z. kombuchaensis* and *Z. lentus*. They share 5 SNPs over the alignment. The most diverse pair of *Zygosaccharomyces* species (*Z. bailii* and *Z. mellis*) share only 46 SNPs. These numbers demonstrate that the D1 / D2 sequence is not a suitable target for primers to discriminate between species of this genus.

The 18S and 5.8S rDNA sequences for the *Zygosaccharomyces* species are even more similar than the D1 / D2 sequence, making them even more unsuitable targets. The 5.8S sequence of *Z. bailii*, *Z. bisporus*, *Z. lentus* and *Z. kombuchaensis* are identical. The 18S sequences show little variation (data not shown).

The ITS sequences are highly polymorphic and seem suitable primer targets. However, because they are not transcribed, the sequences are not constrained. As far as is known, they do not have sequence constraints and so can evolve more freely than the rest of the rDNA. The ITS sequences of multiple strains of *Z. bailii* were elucidated and sequence alignments are illustrated in figures 4.4 and 4.5. Although there are areas of complete homology, there are also areas that have indels between the species. If a primer were designed that encompassed an indel in a previously unsequenced strain of a species, that primer would not be able to anneal to the target. This kind of polymorphism within a primer target would have a more drastic effect on binding than a SNP. The heterogeneity exhibited by the *Z. bailii* strains is illustrated in the phylogenetic trees in figures 4.24 and 4.25. Excluding non-target species would be simple with primers targeted to the ITS, but including all target strains would be a potential problem.

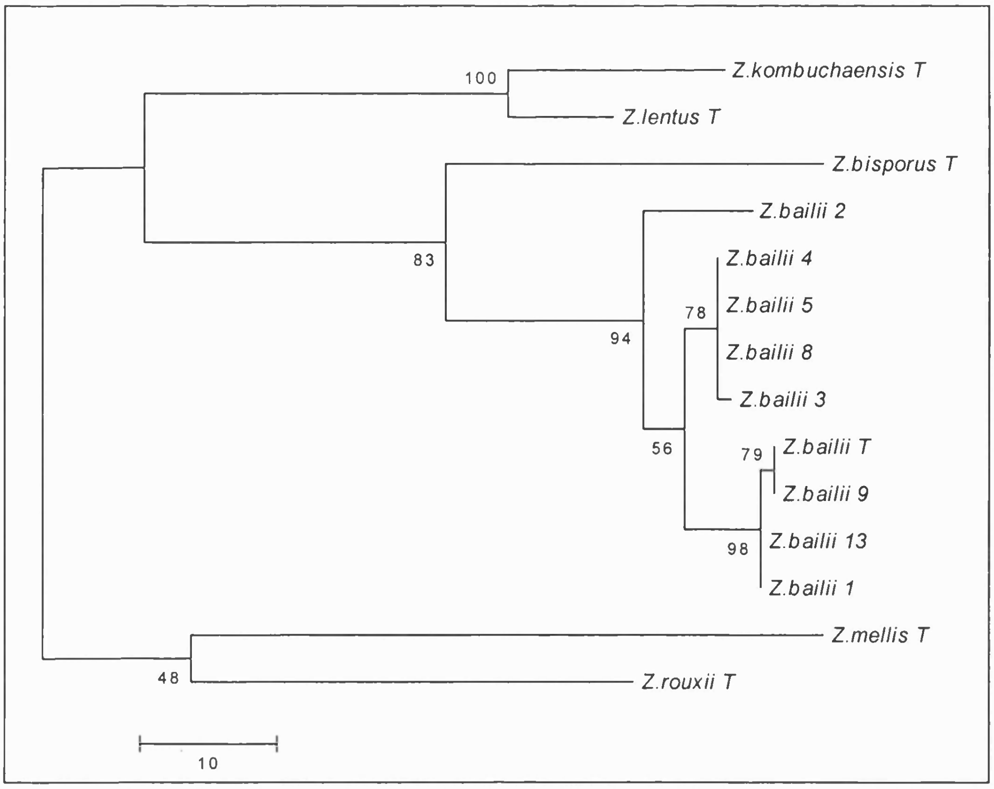

Figure 4.24: A maximum parsimony phylogenetic tree constructed using the ITS1 sequences of the *Zygosaccharomyces* type species and *Z. bailii* strains.

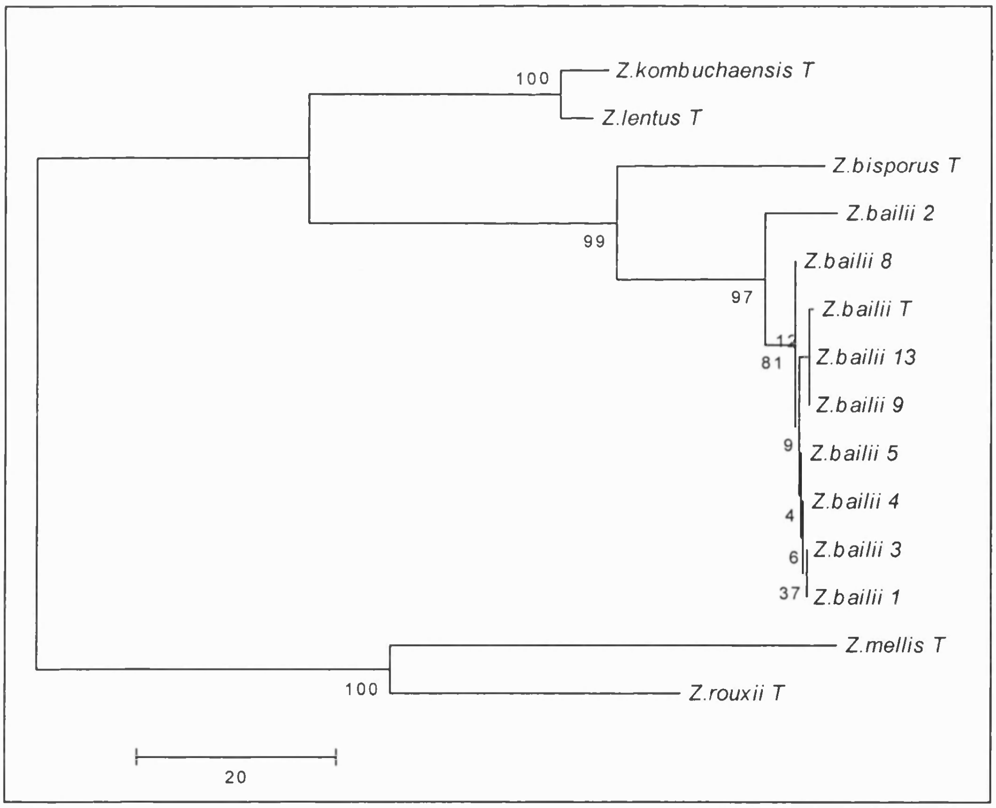

Figure 4.25: A maximum parsimony phylogenetic tree constructed using the ITS2 sequences of the *Zygosaccharomyces* type species and *Z. bailii* strains.

4.3.2 Housekeeping gene sequences

Due to the unsuitability of the rDNA gene sequences for species-specific identification, other areas of the genome were investigated. Housekeeping genes were the next locus investigated. Housekeeping genes are essential to the function of the cell and contained in the genome of all of the yeasts investigated. The fact that they are essential for the viability of the organism means that there are evolutionary constraints upon the protein sequence to stay the same. These constraints are also placed upon the nucleotide sequence, although the degeneracy of the genetic code allows for some polymorphism to occur between even highly related sibling species. The housekeeping genes *HIS3* and *LEU2* were investigated for this genus. The *HIS3* gene was sequenced for all members of the *Zygosaccharomyces* genus. The number of SNPs exhibited between the closest sibling were 43bp (0.08%). That is 38 more than over a greater region of the D1 / D2 sequence and although not ideal, was thought to be adequate for the design of successful species-specific PCR primers. The most distant pairs of species in the *Zygosaccharomyces* genus share 124 SNPs (24%). The closest *Brettanomyces* / *Dekkera* siblings exhibit only 38 SNPs (0.06%) in the D1 / D2 sequence. The *HIS3* gene is therefore a more suitable candidate as a target for specific primers than the rDNA.

The *HIS3* gene was investigated for intraspecific heterogeneity. Multiple *Z. bailii* strains were sequenced. The ClustalW sequence alignment of these strains is shown in figure 4.10 and a phylogenetic tree illustrating the relationships between them is shown in figure 4.26. A phylogenetic tree was constructed using all of the available sequence data available for these strains (figure 4.27) in order to examine their relationships more closely.

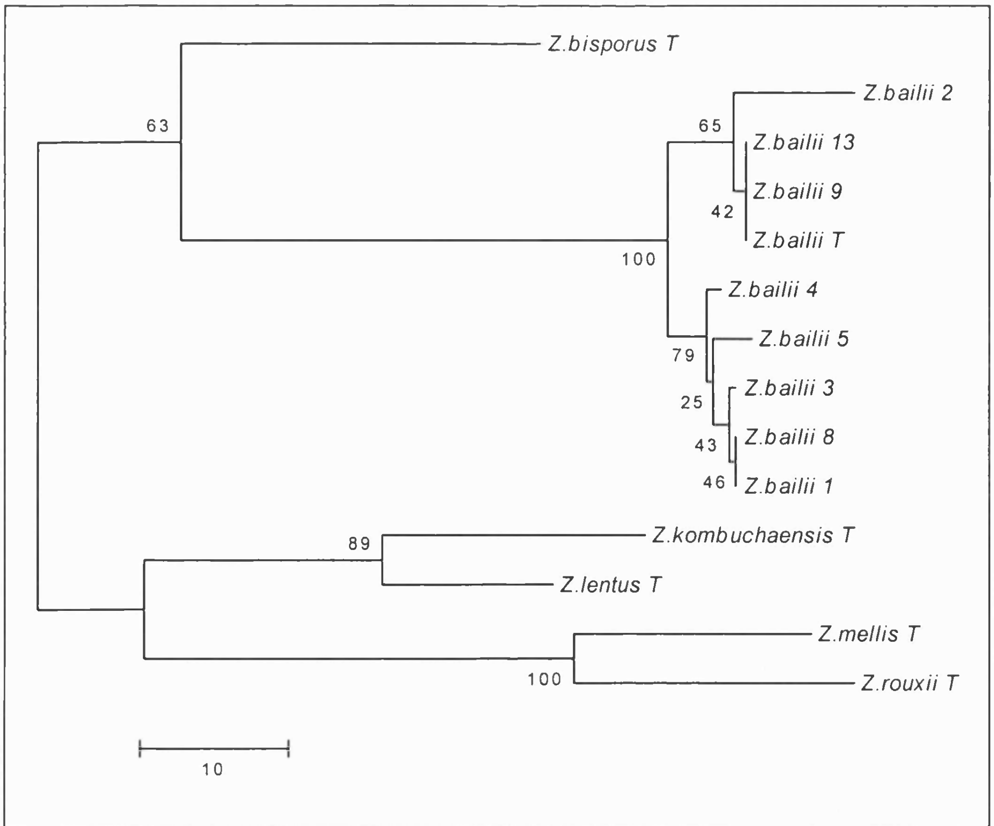

Figure 4.26: A maximum parsimony phylogenetic tree constructed using the *HIS3* sequences of the *Zygosaccharomyces* type species and *Z. bailii* strains.

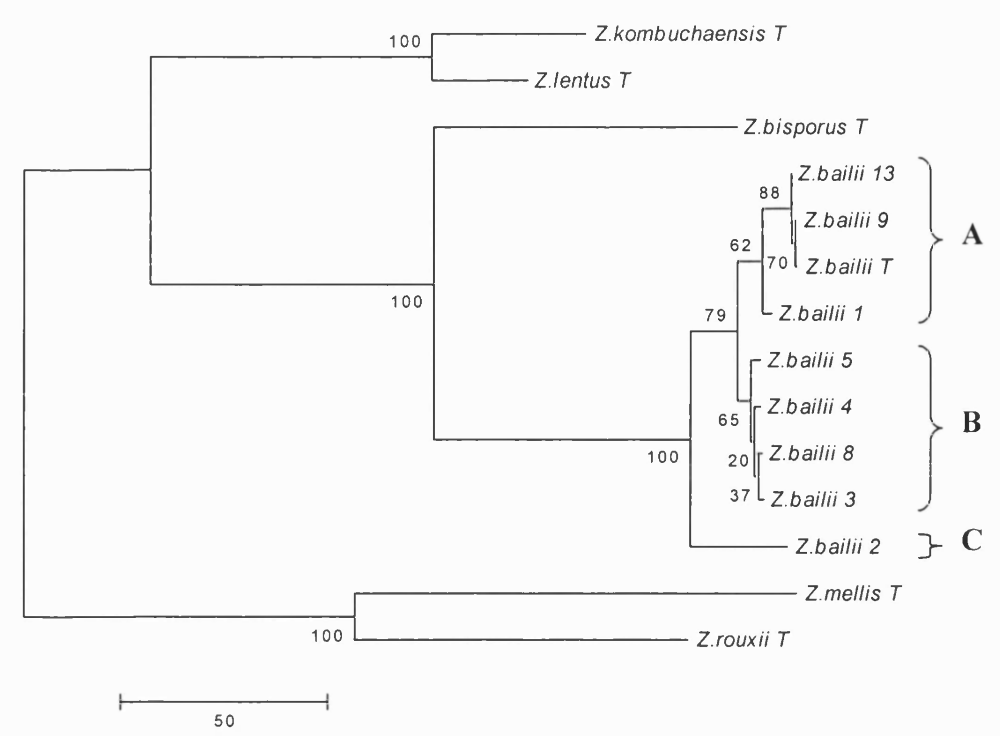

Figure 4.27: A maximum parsimony phylogenetic tree constructed using the ITS1, ITS2 and *HIS3* sequences of the *Zygosaccharomyces* type species and *Z. bailii* strains (*Z. bailii* subtypes A, B and C are marked [28]).

In James *et al* (1996) [28] it was seen that *Z. bailii* ITS sequences separated the taxon into three distinct sequence subtypes. Type A included the type strain (NCYC 1416), NCYC 573 and NCYC 1515. Type B included strains NCYC 1427 (3) and NCYC 1766 (4). Type C included strains NCYC 417 and NCYC 464. The additional strains analysed here also separate into these 3 subtypes. Strains 9 and 13 are subtype A, strains 5 and 8 are subtype B and strain 2 is subtype C. The ITS sequences of strain 2 are almost 100% identical to NCYC 417 (data not shown).

Despite the obvious heterogeneity within the *HIS3* sequence, it was still a more suitable candidate as a target sequence for species-specific PCR primers than any of the rDNA sequences.

4.3.3 Specific primer design

The primary consideration in primer design for the *Zygosaccharomyces* species was to incorporate the maximum possible number of SNP sites between the target and closest sibling into the length of a single primer. As most of the polymorphisms were synonymous substitutions, there were not many consecutive polymorphic sites to include. In order to incorporate as many substitutions as possible, some of the other primer design parameters were broken. Length, base at the 3' position and melting temperature were all secondary considerations compared to polymorphic sites.

When designing specific primers for siblings *Z. lentus* and *Z. kombuchaensis*, the maximum number of differences that could be incorporated into an oligonucleotide was 6. When designing primers for siblings *Z. bailii* and *Z. bisporus*, the maximum number of polymorphisms incorporated into a single oligonucleotide was 9 for *Z. bisporus* and 7 for *Z. bailii*. Siblings *Z. rouxii* and *Z. mellis* had a maximum number of 5 and 8 respectively. These numbers are relatively close as most of the polymorphisms were synonymous substitutions in the third position of a codon. Within a 24-mer oligonucleotide, the maximum number of third position synonymous substitutions is 8.

4.3.4 Specific primer testing

The most important objective of each primer pair is first to show that all target strains are amplified. Despite the heterogeneity within the *HIS3* sequence, the *Z. bailii* species-specific primers amplified all target strains. *Z. bailii* strain 2 was amplified weakly compared to the other strains that were tested (strains 1 – 13 in table 2.3). In a temperature gradient experiment (figure 4.19) the *Z. bailii* 2 amplicon was weakened and then eliminated at a lower temperature than that of strain 1. The *Z. bailii* specific primer loci in some of the *Z. bailii* strains are shown in figure 4.28. Strain 2 has 4 mismatches within the combined target sites. Strains 9 and 13 have 2 mismatches. Strains 4, 5 and 10 have 1 mismatch and strains 1, 3, 8 and 12 have none. Amplification does persist up to an annealing temperature of 69.9°C and so the potential usefulness of the primer pair is not affected. This result shows that although strains can be heterogeneous, the method is still successful in amplifying them all at the same time as eliminating other species.

```
                     10        20
             ....|....|....|....|....
F1           GCCCTAGCGAAGCACTCGGGTTGG
Z.bailii 1   ........................
Z.bailii_2   .....T..A.....T.........
Z.bailii_3   ........................
Z.bailii_4   ........................
Z.bailii_5   ........................
Z.bailii_8   ........................
Z.bailii_9   ..............T.........
Z.bailii_10  ........................
Z.bailii_12  ........................
Z.bailii_13  ..............T.........
                     10        20
             ....|....|....|....|....|
R2           CTTTTCAAACCAAGGTCAATCACCG
Z.bailii_1   .........................
Z.bailii_2   .......................T.
Z.bailii_3   .........................
Z.bailii_4   .......................T.
Z.bailii_5   .......................T.
Z.bailii_8   .........................
Z.bailii_9   .......................T.
Z.bailii_10  .......................T.
Z.bailii_12  .........................
Z.bailii_13  .......................T.
```

Figure 4.28: The *Z. bailii* specific primer target loci in a variety of *Z. bailii* strains.

Once target amplification has been achieved, it is imperative to eliminate the closest sibling species. In some cases, there were enough polymorphism within the target loci to automatically do this. In other cases, after all avenues of primer testing had been explored, the only option remaining was to alter the reaction conditions to eliminate the false positive reactions. The *Z. kombuchaensis* primer pair F3 / R1 amplified no other species except *Z. lentus*. As can be seen from the primer target loci in each *Zygosaccharomyces* species (figure 4.29), the closest sibling *Z. lentus* had 10 mismatches within the primer-binding site. *Z. bisporus* and *Z. bailii* had 14 mismatches. *Z. mellis* had 18 mismatches and *Z. rouxii* had 21. A temperature gradient was performed to eliminate the non-specific binding (figure 4.23). The *Z. lentus* amplicons were eliminated at 66.8°C while the specific amplification persisted up to 69.9°C.

```
                          10        20
                     ....|....|....|....|....|
ZkoF3                CAAAAGGATCACGAATGAGACTAAG
Z.bailii             .C.G.....T..C..C.....C..A
Z.bisporus           ...G..A.....T..C.....C..A
Z.kombuchaensis      .........................
Z.lentus             ...G.....A........A..C...
Z.mellis             .C.G....AT..C..C.....C..A
Z.rouxii             GC.G..A.AT..C..C..A..C..A
                          10        20
                     ....|....|....|....|...
ZkoR1                GCCCAAGTTGATCACTGCGTAAG
Z.bailii             A..A.G..CA.....C..A....
Z.bisporus           ...T.GA.CT..A..A.....G.
Z.kombuchaensis      .......................
Z.lentus             A..T....CA..A........T.
Z.mellis             T.....T.CA..A..A.AAC.T.
Z.rouxii             T..T..T.CA..A..A.AA..T.
```

Figure 4.29: The *Z. kombuchaensis* specific primer target loci in the *Zygosaccharomyces* species.

Once it had been shown that a primer pair amplified all of the target strains and none of the closest sibling strains, it was tested on a variety of strains from the rest of the genus. In most cases, if the sibling had not been amplified, the other strains were not amplified either. This is because the sequence differences in the target loci usually increase as phylogenetic distance increases. However, this was not always the case. The *Z. lentus* primer pair F2 / R1 did not amplify the closest sibling, *Z. kombuchaensis*. They did amplify the majority of the *Z. bisporus* strains and *Z. bailii* strain 7. Non-specific amplification was eliminated between 58.2 – 60.5°C. Specific amplification persisted to 63.5 – 66.8°C (figure 4.17). An annealing temperature of 65°C was found to be *Z. lentus* specific (figure 4.18). Examination of the primer target loci in the *HIS3* gene of each species shows that the nine 3’ nucleotides of the F2 primer are 100% identical in the *Z. bailii* and *Z. bisporus HIS3* target (figure 4.30). The same *Z. kombuchaensis* region has 3 mismatches. Under the generic PCR conditions, the homology at the 3’ end of one primer is enough to result in non-specific amplification. More stringent reaction conditions easily eliminated these amplicons while specific amplification was retained.

```
                            10        20
                    ....|....|....|....|...
ZleF2               CACAGGCCACTGGCTCTCAAGTG
Z.bailii            .G........CACT.........
Z.bisporus          ..........CTCT.........
Z.kombuchaensis     .G..............C..G..C
Z.lentus            .......................
Z.mellis            ....A..T...A.TG.......T
Z.rouxii            .G..A..T...A.T......AC.
                            10        20
                    ....|....|....|....|...
ZleR1               ACCTAAGTCAATAACTGCGTATG
Z.bailii            ...A.G......C..C..A..A.
Z.bisporus          G....GA..T.....A.....G.
Z.kombuchaensis     G..C....TG..C........A.
Z.lentus            .......................
Z.mellis            T..C..T........A.AAC...
Z.rouxii            T.....T........A.AA....
```

Figure 4.30: The *Z. lentus* specific primer target loci in the *Zygosaccharomyces* species. Homologous regions in the *Z. bailii* and *Z. bisporus* sequence and the corresponding non-homologous *Z. kombuchaensis* region are underlined.

Once all *Zygosaccharomyces* species had been eliminated, outgroup species were tested. It would be reasonable to assume that the sequence of the *HIS3* gene would be increasingly dissimilar to the target species as phylogenetic distance increased outside of the genus. A phylogenetic tree of the *Zygosaccharomyces* species and the outgroup species tested is shown in figure 4.31.

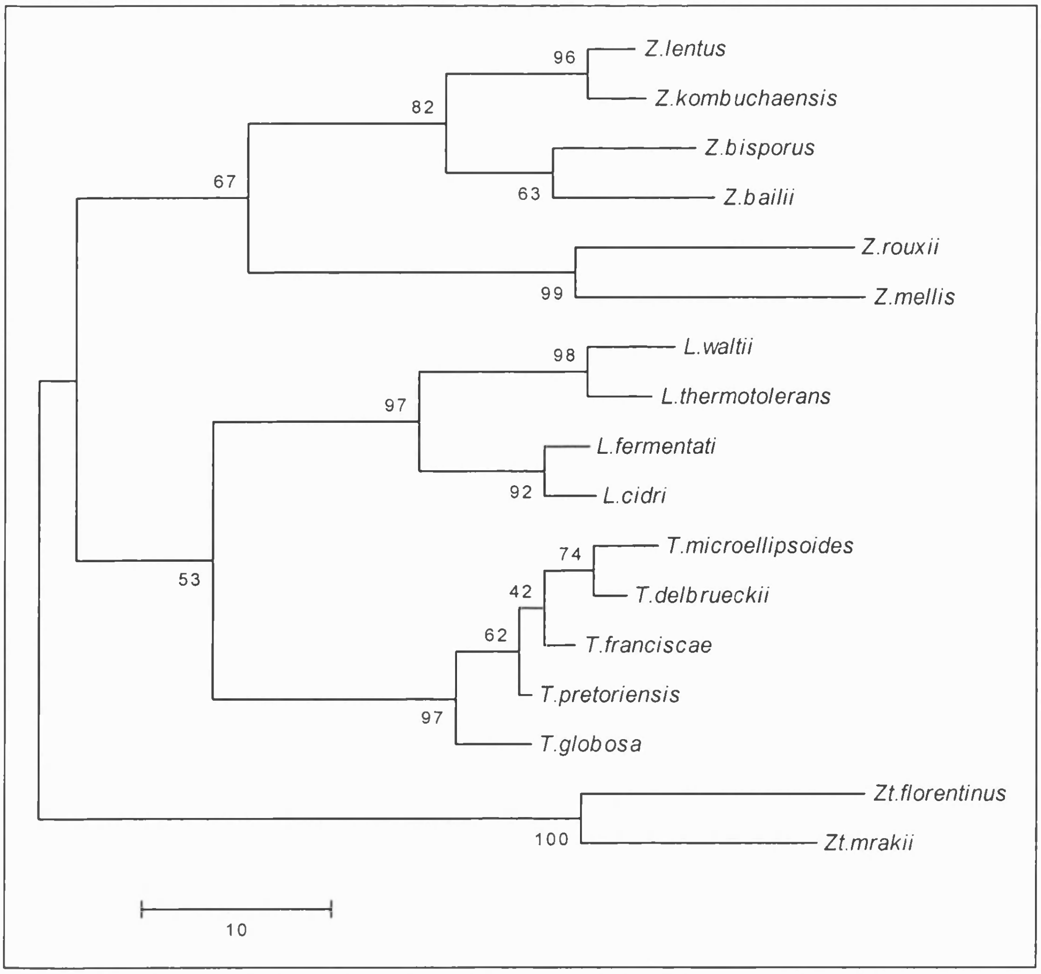

Figure 4.31: A maximum parsimony phylogenetic tree built using the D1 / D2 sequences of the type strains of clades 7, 8, 9 and 10 of the *Saccharomycetaceae*. NCBI accession numbers listed in table 2.5).

The *Z. bailii* primer pairs F3 / R1 and F3 / R2 annealed and amplified the target locus in *L. cidri* despite not amplifying any *Zygosaccharomyces* barring the target species. *Z. bailii* primer pair F1 / R2 did not amplify the *L. cidri HIS3* gene. Examination of the primer binding loci in the *L. cidri HIS3* gene (figure 4.32) revealed that the F3 primer had higher homology to the *L. cidri* target than it did to *Z. bisporus*, the closest sibling of *Z. bailii*. A ClustalW alignment of the partial *HIS3* sequences of the *Zygosaccharomyces* species and *L. cidri* (figure 4.16) shows that this portion of the *HIS3* gene is not representative of the whole sequence and *L. cidri* shares considerable sequence diversity with *Z. bailii*. The similarity of the F3 target sequence in these two species is merely coincidence.

```
Z.bailii_F1        GCCCTAGCGAAGCACTCGGGTTGG
Z.bailii           ........................
L.cidri            ------..T.....T..A..C...
Z.bisporus         ..A........A.....C..G...
Z.lentus           ..G..............C......
Z.kombuchaensis    ..G..C...........C......
Z.rouxii           ..T..G.....A.....T......
Z.mellis           ..T..G.....A.....A......

Z.bailii_F3        CTGAGGATTGTGGCATTGCTCTCG
Z.bailii           ........................
L.cidri            .C...........T..........
Z.bisporus         .A.....C.....T.....G..A.
Z.lentus           ....A..C.....T.....C..T.
Z.kombuchaensis    ....A..C.....T.....C..T.
Z.rouxii           ....A..C..C..T.....C..A.
Z.mellis           ....A..C..C..T.....C..A.

Z.bailii_R1        TAGACAAGTCAACGACAGCACGCG
Z.bailii           ........................
L.cidri            .G..T..A..C..A..C.......
Z.bisporus         .G....GA..C.............
Z.lentus           .G.....A..G.....T.....G.
Z.kombuchaensis    .G.....A..G.....G.......
Z.rouxii           .G..T.GA..G..A.....T.TT.
Z.mellis           .......A..G..T.....T.TT.

Z.bailii_R2       CTTTTCAAACCAAGGTCAATCACCG
Z.bailii          .........................
L.cidri           ..C.....T..C...........T.
Z.bisporus        ........G..T..A..T..A..A.
Z.lentus          ...........T.A......A..T.
Z.kombuchaensis   ........G..C.A..TG.....T.
Z.rouxii          ........T..T.AT.....A..A.
Z.mellis          .....T..T..C.AT.....A..A.
```

Figure 4.32: *Z. bailii* primer target regions in *Zygosaccharomyces sp.* and *L. cidri*. F3 / R1 and F3 / R2 are positive with *L. cidri*. The high homology of the F3 primer with *L. cidri* is highlighted

This reaction had implications on the *Z. bailii* specific primer pair. The primer pairs that amplified *L. cidri* strains were the only primers that did not amplify any of the other *Zygosaccharomyces* species. The *L. cidri* amplification had been unexpected and suggested that the F3 primer target was a conserved region in the *HIS3* gene. For this reason, the *Z. bailii* F1 / R2 primer pair that had previously amplified *Z. lentus* strains and *Z. rouxii* strain 3, underwent reaction condition optimisation (figure 4.19). A temperature gradient eliminated *Z. lentus* at 68.4 – 69.9°C and *Z. rouxii* at 55.5 – 58.4°C. *Z. rouxii* amplification was very weak and inconsistent throughout. The heterogeneity of the *Z. bailii* strains has been discussed previously. *Z. bailii* strain 2 of subgroup C is the most atypical *Z. bailii* tested and amplification was eliminated at 69.9 - 70°C while strain 1 persisted to 70°C. As the *Z. lentus* strains were eliminated at a similar temperature as *Z. bailii* strain 2 was lost, the higher temperature could not be used in a specific reaction and still be guaranteed to amplify *Z. bailii* type C species.

A magnesium ion gradient at 55°C eliminated *Z. lentus* amplification at 2.5 – 2.0mM Mg^{2+}. *Z. bailii* strain 1 amplification persisted at 1.0mM Mg^{2+} and *Z. bailii* strain 2 amplification persisted to 1.6mM – 1.5mM Mg^{2+} (figures 4.20 and 4.21). A combination of temperature and magnesium ion optimisation resulted in the *Z. bailii* F1 / R2 primers becoming specific at 60°C and 1.8mM Mg^{2+}. Subtypes A, B and C were amplified under these reaction conditions and non-specific amplification was eliminated (figure 4.22).

4.4 Conclusion

The *HIS3* housekeeping gene provided a suitable target for species-specific PCR identification of this monophyletic, closely related genus. Species that are notoriously difficult to distinguish between using physiological methods were easily differentiated. The *Z. mellis* type strain is difficult to distinguish from *Z. rouxii* using chemotaxonomy [127]. However, the limitations imposed by only having approximately 516bp of sequence in which to find a sufficient amount of polymorphism to distinguish between 6 species were considerable. There was a limit to primer choice that resulted in some less than ideal primers having to be optimised by temperature and magnesium ion alterations in order to be specific. Although this is a common problem with PCR, the different conditions for each primer pair have practical implications for a future high throughput method. Separate PCR blocks would be required to test the *Z. bailii*, *Z. lentus* and *Z. kombuchaensis* primers while *Z. rouxii* and *Z. mellis* primers are specific under the same conditions. Additionally, the different magnesium ion concentration required for *Z. bailii* specificity would require a different reaction mixture than the others. These obstacles are not insurmountable, but they are not ideal. The different reaction conditions also mean that a multiplex reaction with all primer pairs in one reaction mixture would be impossible.

The possible discovery of novel *Zygosaccharomyces* species and hybrids has been recently reported [134]. The impact of new species on the species-specific PCR primer method would not be particularly grave. The *HIS3* gene could be sequenced from the new species and primers could be designed and tested in the same way as before. However, the current specific primers would also need to be tested to ensure that reactions were negative. Depending on how closely a novel species was related to the current species would determine the likelihood the method should still be valid.

Hybrids are a different issue. The success or failure of the current primers to amplify a hybrid species would depend on the parental species of the hybrid and the nature of the *HIS3* gene. If the parental species of a hybrid were *Z. rouxii* and *Z. mellis* and the hybrid retained a copy of each parental *HIS3* gene, it would be identified by both the *Z. rouxii* and *Z. mellis* specific primers. This would be a clear indication that a hybrid was present. If the hybrid retained only one of the parental copies, it would be misidentified as that parent. On the other hand, a hybrid copy or copies of the gene could have since diverged and then it would not be amplified with any of the *Zygosaccharomyces* primers. This would give the false impression that the strain being investigated was a non-*Zygosaccharomyces* yeast. Extensive sequence surveys of any hybrids would be required to determine their origins and the nature of their *HIS3* genes before conclusions could be drawn. However, the fact that natural yeast hybrids seem to be more prevalent than first thought [134, 135, 136] may have future implications for this method.

Chapter 5. *Saccharomyces*

5.1 Introduction

A description and history of this target genus is described below with a brief explanation of spoilage characteristics.

5.1.1 Currently accepted species

Saccharomyces bayanus Saccardo (1895)

Saccharomyces cariocanus Naumov, James, Naumova, Louis & Roberts (2000)

Saccharomyces cerevisiae Meyen ex E. C. Hansen (1883)

Saccharomyces kudriavzevii Naumov, James, Naumova, Louis & Roberts (2000)

Saccharomyces mikatae Naumov, James, Naumova, Louis & Roberts (2000)

Saccharomyces paradoxus Bachinskaya (1914)

Saccharomyces pastorianus E. C. Hansen (1904)

5.1.2 Description of the genus

Vegetative reproduction - Asexual reproduction is by multilateral budding on a narrow base. Cells are spheroidal, ovoidal or elongate. Pseudohyphae may be formed, but true hyphae are not produced.

Ascospore formation - Asci are unconjugated, persistent and produce 1 – 4 spherical to ovoidal ascospores.

Physiology / biochemistry – Glucose, raffinose and usually sucrose are fermented, often vigorously. Cadaverine, L-lysine, ethylamine and nitrate are not utilized as nitrogen sources. Coenzyme Q-6 is produced. The diazonium blue B reaction is negative [137].

5.1.3 Spoilage characteristics

The appearance of *Saccharomyces* species as spoilage organisms in soft drinks often indicates yeasts escaped from wineries, breweries or bakeries. They are the main contaminant in installations producing both beer and fruit juices. There are many strains and variants that show considerable variation in resistance to preservatives. Characteristically highly fermentative yeasts, *Saccharomyces* metabolise glucose, fructose and sucrose with ease and are fast growers. Spoilage results in alcoholic fermentation, pressure build-up and clouds and haze [107].

5.1.4 History

The *Saccharomyces sensu stricto* group is currently composed of the species *S. cerevisiae*, *S. bayanus*, *S. paradoxus*, *S. pastorianus*, *S. mikatae*, *S. kudriavzevii* and *S. cariocanus*. The species *S. cerevisiae*, *S. bayanus* and *S. pastorianus* have been associated with brewing, wine making and baking for millennia. They are the most investigated yeast species in existence and have been exploited in man-created environments for centuries. Many strains investigated have been found to be aneuploid or polyploid. *S. cerevisiae* is a paleopolyploid. After a whole genome duplication event that diverged the *Saccharomyces* genus from *Kluyveromyces* [138, 139, 140] it went through a polyploid phase until the chromosomes became distinct again. The loss of many of the duplicated genes and asymmetric evolution of the retained duplicate pairs [139, 141, 142] has added to the complexity of the genome.

S. pastorianus is thought to have arisen as a hybrid of *S. cerevisiae* and *S. bayanus* [143]. The type strain (CBS 1538) has an identical D1 / D2 sequence to *S. bayanus* and yet is physiologically more similar to *S. cerevisiae*. *S. cerevisiae* and *S. bayanus* have been shown to occasionally mate and produce hybrids with a low level of fertility [144].

There are also many strains that are now classified within the umbrella of *S. cerevisiae* synonyms. Despite the *Saccharomyces* genus being well studied, new species are still being found. Three new *sensu stricto* species were announced in 2000: *S. cariocanus*, *S. mikatae* and *S. kudriavzevii* [145].

S. pastorianus strains are allopolyploid, containing parts of two divergent genomes. They are sterile and so genetic investigations are limited. The parental species or strains that contributed to the original hybrid, or hybrids are the subject of much investigation. It has long been established that *S. cerevisiae* and a member, or members of the *S. bayanus* group are the parent species [146]. The *S. pastorianus* strains form a heterogeneous taxon. *S. bayanus* strains have also been reported to be heterogeneous [147, 148], they are split into two subgroups: a "*bayanus*" subgroup including the *S. bayanus* type strain (CBS 380) and a "*uvarum*" subgroup including the *S. uvarum* synonym type strain (CBS 395) and the sequence surveyed *S. bayanus* strain MCYC 623 (CBS 7001).

The "*uvarum*" subgroup within *S. bayanus* species has been shown to form a highly homogeneous group, prompting calls for classification as a discrete taxon [149]. The remaining *S. bayanus* strains are not homogeneous. It has been proposed that the type strain (CBS 380) is a hybrid of the *S. uvarum* type strain and *S. cerevisiae* [150]. However, as it is likely that a member of this heterogeneous group was the other parent of the heterogeneous hybrid strains, the origins of the current group of strains that are classified as *S. pastorianus* are difficult to ascertain.

The former *S. carlsbergensis* and *S. monacensis* type strains (CBS 1513 and CBS 1503 respectively) are classified as *S. pastorianus.* Lodder and Kreger-van Rij (1952) [151] hypothesised that *S. cerevisiae* and *S. monacensis* were the parental species of *S. pastorianus.* CBS 1503 and CBS 1513 have both since been shown to be hybrids [152]. Investigations of various genetic markers in *S. pastorianus* CBS 1503, CBS 1513 and CBS 1538 have shown that each markers origin varies within single strains. They originate from a combination of sources derived from *S. bayanus* subgroup "*bayanus*", *S. bayanus* subgroup "*uvarum*" and *S. cerevisiae.* Some of the markers are 100% identical to one of these groups; others are slightly diverged from the ancestral sequence but exhibit an obvious lineage [152, 153].

It seems reasonable to hypothesise that such a heterogeneous group will be difficult to identify using a single pair of PCR primers. A single locus that is identical in all *S. pastorianus* strains will be difficult to find without extensive genome surveys. Even then, it may be impossible to find if they have arisen from multiple hybrid events as opposed to a single rare mating.

5.2 Results

5.2.1 Target sequences

From the *Saccharomyces sensu stricto* (clade 1 [133]) D1 / D2 alignment in figure 5.1 it is obvious that these species are as closely related as the *Zygosaccharomyces* genus and the rDNA is therefore not suitable for PCR species identification. However, complete genome sequence is available for *S. cerevisiae* and extensive genome sampling has been performed for *S. bayanus*, *S. mikatae* and *S. paradoxus.* The choice of gene sequence data was therefore not limited. A different single gene was used as a primer target for each individual species.

In a study that attempted to find the minimum number of genes required for a completely resolved phylogenetic tree [35], a list was compiled of *Saccharomyces* genes that did not have any orthologues within any *Saccharomyces* genome. Genes were chosen from this list that were lethal when knocked out in *S. cerevisiae*, therefore having more evolutionary pressure to stay the same. These genes are listed in table 5.1. The redundancy of the genetic code would result in enough variability between species. Nine genes were chosen (after an initial alignment to look for adequate degrees of variation) for the seven species to allow for some to be unsuitable. These genes were *DBP6*, *FAL1*, *MEX67*, *PRI1*, *SEC24*, *SEC31*, *SPB1*, *SRP72* and *VAS1*. A list of the systematic names and functions of these genes is shown in table 5.2. The homologue sequences of *S. bayanus*, *S. cerevisiae*, *S mikatae* and *S. paradoxus* were located and downloaded using the genome wide alignment synteny viewer tool on the *Saccharomyces* genome database [154] based on Kellis *et al* (2003) [155]. Some sequences had to be experimentally obtained for *S. pastorianus*, *S. cariocanus* and *S. kudriavzevii.* Partial *S. kudriavzevii* genome sequences for all genes except *MEX67* were available from The Washington University *Saccharomyces* genome website [156, 157]. *S. pastorianus* gene sequences were obtained from Dr. Kodama at Suntory Research Centre (Personal Communication). *S. cariocanus* sequences were found after designing degenerate primers from the alignment of *S. bayanus*, *S. cerevisiae*, *S. mikatae* and *S. paradoxus* genes. A similar approach was tried for *S. kudriavzevii MEX67* gene. Amplicons were inconsistently produced and impossible to sequence. The work was continued without this sequence.

```
                                   10        20        30        40        50        60        70        80        90       100
                         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.cerevisiae_NCYC_505    AGCGGCAAAAGCTCAAATTTGAAATCTGGTACCTTCGGTGCCCGAGTTGTAATTTGGAGAGGGCAACTTTGGGGCCGTTCCTTGTCTATGTTCCTTGGAA
S.bayanus_NCYC_2578      ...................................T.....................................A..........................
S.paradoxus_NCYC_2600    ....................................................................................................
S.mikatae_NCYC_2888      ...................................T..............................T..C.A.A..........................
S.kudriavzevii_NCYC_2889 ........................................................................AA..........................
S.pastorianus_NCYC_392   ...................................T.....................................A..........................
S.cariocanus_NCYC_2890   ....................................................................................................
                                  110       120       130       140       150       160       170       180       190       200
                         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.cerevisiae_NCYC_505    CAGGACGTCATAGAGGGTGAGAATCCCGTGTGGCGAGGAGTGCGGTTCTTT--GTAAAGTGCCTTCGAAGAGTCGAGTTGTTTGGGAATGCAGCTCTAAG
S.bayanus_NCYC_2578      .................................................A.--...............................................
S.paradoxus_NCYC_2600    .................................................A.--...............................................
S.mikatae_NCYC_2888      .................................................CAAT..G..A.........................................
S.kudriavzevii_NCYC_2889 ...................................................T-...............................................
S.pastorianus_NCYC_392   .................................................A.--...............................................
S.cariocanus_NCYC_2890   ...................................................--...............................................
                                  210       220       230       240       250       260       270       280       290       300
                         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.cerevisiae_NCYC_505    TGGGTGGTAAATTCCATCTAAAGCTAAATATTGGCGAGAGACCGATAGCGAACAAGTACAGTGATGGAAAGATGAAAAGAACTTTGAAAAGAGAGTGAAA
S.bayanus_NCYC_2578      ....................................................................................................
S.paradoxus_NCYC_2600    ....................................................................................................
S.mikatae_NCYC_2888      ....................................................................................................
S.kudriavzevii_NCYC_2889 ....................................................................................................
S.pastorianus_NCYC_392   ....................................................................................................
S.cariocanus_NCYC_2890   ....................................................................................................
                                  310       320       330       340       350       360       370       380       390       400
                         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.cerevisiae_NCYC_505    AAGTACGTGAAATTGTTGAAAGGGAAGGGCATTTGATCAGACATGGTGTTTTGTGCCCTCTGCTCCTTGTGGGTAGGGGAATCTCGCATTTCACTGGGCC
S.bayanus_NCYC_2578      .....................................................C....................G.............G...........
S.paradoxus_NCYC_2600    ....................................................................................................
S.mikatae_NCYC_2888      ....................................................................................................
S.kudriavzevii_NCYC_2889 .....................................................C....................G.........................
S.pastorianus_NCYC_392   .....................................................C....................G.............G...........
S.cariocanus_NCYC_2890   ....................................................................................................
                                  410       420       430       440       450       460       470       480       490       500
                         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.cerevisiae_NCYC_505    AGCATCAGTTTTGGTGGCAGGATAAATCCATAGGAATGTAGCTTGCCTCGGTAAGTATTATAGCCTGTGGGAATACTGCCAGCTGGGACTGAGGACTGCG
S.bayanus_NCYC_2578      .............................G..........A.....T....G..............A..............A..................
S.paradoxus_NCYC_2600    .............................G..........A.....T....G...............C................................
S.mikatae_NCYC_2888      .............................G..........A.....T....G...............C................................
S.kudriavzevii_NCYC_2889 .............................G..........A.....T....G...............C................................
S.pastorianus_NCYC_392   .............................G..........A.....T....G..............A..............A..................
S.cariocanus_NCYC_2890   .............................G..........A.....T....G...............C................................
```

Figure 5.1: ClustalW alignment of the *Saccharomyces sensu stricto* type species D1 / D2 rDNA region. NCBI accession numbers listed in table 2.5

Genes were then matched for suitability to a species, in some cases depending upon species-specific motifs in the sequence. Others were arbitrarily assigned.

Systematic name	Standard name	Sequences unavailable
YAR007C	*RFA1*	*S.cariocanus, S.pastorianus*
YBR110W	*ALG1*	*S.cariocanus, S.pastorianus*
YBR198C	*TAF5*	*S.cariocanus, S.pastorianus*
YCL054W	*SPB1*	*S.cariocanus, S.pastorianus*
YDL031W	*DBP10*	*S.cariocanus, S.kudriavzevii, S.pastorianus*
YDL126C	*CDC48*	*S.cariocanus, S.pastorianus*
YDL148C	*NOP14*	*S.cariocanus, S.kudriavzevii, S.pastorianus*
YDL166C	*FAP7*	*S.cariocanus, S.kudriavzevii, S.pastorianus*
YDL195W	*SEC31*	*S.cariocanus, S.pastorianus*
YDR021W	*FAL1*	*S.cariocanus, S.pastorianus*
YDR054C	*CDC34*	*S.cariocanus, S.kudriavzevii, S.pastorianus*
YDR361C	*BCP1*	*S.cariocanus, S.pastorianus*
YGL001C	*ERG26*	*S.cariocanus, S.pastorianus*
YGL225W	*VRG4*	*S.cariocanus, S.pastorianus*
YGR005C	*TFG2*	*S.cariocanus, S.pastorianus*
YGR094W	*VAS1*	*S.cariocanus, S.pastorianus*
YHR019C	*DED81*	*S.cariocanus, S.kudriavzevii, S.pastorianus*
YIL109C	*SEC24*	*S.cariocanus, S.pastorianus*
YIR008C	*PRI1*	*S.cariocanus, S.pastorianus*
YJL085W	*EXO70*	*S.cariocanus, S.pastorianus*
YJL087C	*TRL1*	*S.cariocanus, S.kudriavzevii, S.pastorianus*
YJR068W	*RFC2*	*S.cariocanus, S.pastorianus*
YKL104C	*GFA1*	*S.cariocanus, S.pastorianus*
YKR071C	*DRE2*	*S.cariocanus, S.pastorianus*
YMR203W	*TOM40*	*S.cariocanus, S.pastorianus*
YMR277W	*FCP1*	*S.cariocanus, S.kudriavzevii, S.pastorianus*
YNL062C	*GCD10*	*S.cariocanus, S.pastorianus*
YNL287W	*SEC21*	*S.cariocanus, S.pastorianus*
YNR038W	*DBP6*	*S.cariocanus, S.pastorianus*
YOL097C	*WRS1*	*S.cariocanus, S.kudriavzevii, S.pastorianus*
YOR361C	*PRT1*	*S.cariocanus, S.pastorianus*
YPL028W	ERG10	*S.kudriavzevii, S.cariocanus, S.pastorianus*
YPL169C	MEX67	*S.kudriavzevii, S.cariocanus*
YPL210C	SRP72	*S.cariocanus, S.pastorianus*
YPR181C	SEC23	*S.kudriavzevii, S.cariocanus, S.pastorianus*

Table 5.1: *Saccharomyces* lethal knockout genes

Standard name	**Systematic name**	**Description**
SPB1	YCL054W	Suppressor of PaB1 mutant; involved in 60S ribosomal subunit biogenesis
SEC31	YDL195W	Involved in protein transport from endoplasmic reticulum to Golgi
FAL1	YDR021W	Nucleolar protein required for maturation of 18S rRNA, member of the eIF4A subfamily of DEAD-box ATP-dependent RNA helicases
VAS1	YGR094W	Mitochondrial and cytoplasmic valyl-tRNA synthetase
SEC24	YIL109C	The Sec23p-*SEC24*p complex is one of three cytoplasmic COPII factors involved in ER to Golgi transport
PRI1	YIR008C	p48 polypeptide of DNA primase
DBP6	YNR038W	Dead Box Protein 6
MEX67	YPL169C	Involved in nuclear mRNA export, binds both poly(A) tail and nuclear pore
SRP72	YPL210C	Part of the signal recognition particle (SRP) ribonucleoprotein (RNP) complex that functions in protein targeting to the endoplasmic reticulum (ER) membrane

Table 5.2: *Saccharomyces* target genes, systematic names, functional names and brief description of their function.

5.2.2 Specific primer design

Potentially species-specific primers were designed in the same manner and using the same criteria as those for *Brettanomyces / Dekkera* species (section 3.4.2). All *Saccharomyces* potentially specific primers are listed in table 5.3 with amplicon sizes in table 5.4.

Saccharomyces	**Primer sequence, 5' to 3'**	**Tm (°C)**	G.C (%)	**Length (bp)**
S.bayanus DBP6				
Sbay F1	GCT GAC TGC TGC TGC TGC CCC CG	67	74	23
Sbay F2	AAG GCT CCG ATA ATG ATG ACC GT	57	48	23
Sbay R1	TGT TAT GAG TAC TTG GTT TGT CG	52	39	23
Sbay R2	ACG GTG ATC ATT ATC GGA GCC TT	57	48	23
NB: F2 and R2 are the same sequence but in different directions, therefore the only possible combinations are F1/R1, F1/R2 and F2/R1.				
S.cariocanus VAS1				
Scar F1	GGG CTT AAA ATT AAC TGG TTC AA	50	35	23
Scar F2	GAC GAT TAT CAA CCA CCT ACT AT	51	39	23
Scar F3	GTG ATG TCT ATA TCG AGA ACT CT	52	39	23
Scar R1	CAC AAC CTT CAG GAA TTT CAG AG	53	43	23
Scar R2	ATC TTT CTG ATC CTT AGC GGT T	53	41	22
Scar R3	AGA GGC CTT TAC AAT TGC GGC G	59	55	22
S.cerevisiae MEX67				
Scer F1	TAG ATA ACA ATG GGG CCT CTG CG	58	52	23
Scer F2	GCG CTT TAC ATT CAG ATC CCG AG	58	52	23
Scer F3	TCA ATC TTG CTG ACA ACC AAC TT	54	39	23
Scer R1	GTG CGC CGG GAA TGC TCA TCC CT	63	65	23
Scer R2	TAA GTT GGT TGT CAG CAA GAT TG	53	39	23
NB: F3 and R2 are the same sequence but in different directions, therefore F3/R2 combination is not possible.				
S.kudriavzevii PRI1				
Skud F1	TTC AAG GCA CAG ATA GAG AGA AGT	56	42	24
Skud F2	ATC TAT AAC AAA CCG CCA AGG GAG	56	46	24
Skud F3	ACG GAA GAT CTT GGT TAC AAA GAT	53	38	24
Skud R1	CGT AAC CTA CCT ATA TGA GGG CCT	56	50	24
Skud R2	TCT AAC CAA TGC TTT CTC AAT GAG	54	38	24
S.mikatae FAL1				
Smik F1	ACA AGC AAT TGA TTT GAG GAA AAG	53	33	24
Smik F2	AGG GAT TTG CAA GCT TTG ATA TTG	55	38	24
Smik R1	CCA GTC TTC TTT GTC AAC GTT G	53	45	22
Smik R2	CAC GAG AAT GAC CTG TCC TAA AT	54	43	23

Table 5.3: Sequences of potentially *Saccharomyces* species-specific primers (*continued on next page*).

Saccharomyces	Primer sequence, 5' to 3'	Tm (°C)	G.C (%)	Length (bp)
S.paradoxus SEC31				
Spar F1	AAG CTA AGT ATA CAA ACT CTA AGG	51	33	24
Spar F2	CTT CTA CCA CAT CTA TCC CTA GTG	54	46	24
Spar R1	GGG CCA GCG GGA GGC TTG TTG	61	71	21
Spar R2	AAC CTG GTC GAG CTA CTT GTT TTG	57	46	24
S.pastorianus SPB1				
Spas F1	TCA AGA AAT ACA GCA GCA AGA ACC	56	42	24
Spas F2	AAG AAA TAC AGC AGC AAG AAC CAG	56	42	24
Spas F3	AAA TAC AGC AGC AAG AAC CAG AAG	56	42	24
Spas R1	AAA CGA GCC ACA GAA CGC ATC CGT	62	54	24
Spas R2	TTT TCT GTT TTG TCA GAG TCA TCA	55	38	24

Table 5.3 (*continued*): Sequences of potentially *Saccharomyces* species-specific primers.

		Expected amplicon sizes for each Saccharomyces specific primer pair tested		
Species	Primer	R1	R2	R3
S. bayanus	F1	275	223	
	F2	128		
S. cariocanus	F1	1016	940	760
	F2	638	552	612
	F3	456	383	203
S. cerevisiae	F1	1299	252	
	F2	1197	150	
	F3	1071		
S. kudriavzevii	F1	720	590	
	F2	660	539	
	F3	420	347	
S. mikatae	F1	508	707	
	F2	486	685	
S. paradoxus	F1	1056	265	
	F2	996	202	
S. pastorianus	F1	465	522	
	F2	463	520	
	F3	460	517	

Table 5.4: Expected amplicon sizes for all *Saccharomyces* potentially specific primer pairs.

```
                           10        20        30        40        50        60        70        80        90       100
                  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.mikatae         MSFNRDEDLKLKFKTSKKLKVSPTFESMNLKDDLLRGIYSYGFDAPSSIQSRAITQIISGKDVIAQAQSGTGKTATFTIGLLQAIDLRKRDLQALILSPT
S.cerevisiae      ...D.E..Q.............S....................E..............................................K..........
S.bayanus         .....E..Q............TS.....D..E...........E............................................K.K..........
S.paradoxus       .....E..Q.............SN....K..............E..............................................K..........
S.kudriavzevii    .....E..QR............S....................E............................................KRM..........
S.pastorianus     .....E..Q............TS.....D..E...........E..............................................K..........
S.cariocanus      -------------------...S....................E..............................................K..........
                          110       120       130       140       150       160       170       180       190       200
                  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.mikatae         RELASQIGQVVKNLGDYMNVNAFAMTGGKTLKDDLKKVQKHGCQVVSGTPGRVLDMIKRQMLQTRNVQMLVLDEADELLSETLGFKQQIYDIFAKLPKNC
S.cerevisiae      ........................I............M......A.............K.........................................
S.bayanus         ....N......T.........................L..N...........................V.I.............................
S.paradoxus       .....................................M..............................................................
S.kudriavzevii    ...........T.........................I..N....I.....................R..I......................T......
S.pastorianus     ....N......T.........................L..N...........................V.I.............................
S.cariocanus      .....................................I..............................................................
                          210       220       230       240       250       260       270       280       290       300
                  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.mikatae         QVVVVSATMNKDILEVTRKFMNDPVKILVKRDEISLEGIKQYVVNVDKEDWKFDTLCDIYDSLTITQCVIFCNTKKKVDWLSQRLSQSNFAVVSMHGDMK
S.cerevisiae      .................................................E....................................I..............
S.bayanus         .....................................................................................T..............
S.paradoxus       ..........................................I..........................................L..............
S.kudriavzevii    ....................L.....................I..........................................T..............
S.pastorianus     .....................................................................................T..............
S.cariocanus      .....................................................................................L..............
                          310       320       330       340       350       360       370       380       390       400
                  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.mikatae         QEERDKVMNEFRTGHSRVLISTDVWARGIDVQQVSLVINYDLPEIIENYIHRIGRSGRFGRKGVAINFITKTDSAKLREIEKFYSIKIKPMPANFAELS*
S.cerevisiae      .........D.............................................................A.L..............N..........*
S.bayanus         .........D.....................................................................Q..V...............*
S.paradoxus       .........D.............................................................A................NL.........*
S.kudriavzevii    .........D............................................................RA..T..............L.........*
S.pastorianus     .........D.............................................................A.........Q..V...............*
S.cariocanus      .........D...........................................-----------------------------------------------
```

Figure 5.2: ClustalW amino acid alignment of the *Saccharomyces sensu stricto* species *FAL1* gene. *S. mikatae* tested primer loci are highlighted.

5.2.3 Specific primer testing

S. mikatae

S. mikatae primers were targeted to the *FAL1* gene. A ClustalW amino acid alignment of this gene from all *Saccharomyces sensu stricto* species is shown in figure 5.2. Figure 5.3 shows the relationship between *Saccharomyces* species in this gene sequence using a maximum parsimony phylogenetic tree of the genus. Two forward and two reverse primers were designed, sequences are listed in table 5.3 and shown in figure 5.4 with their target sites in each *Saccharomyces* species. There were four combinations of potentially specific primer pairs.

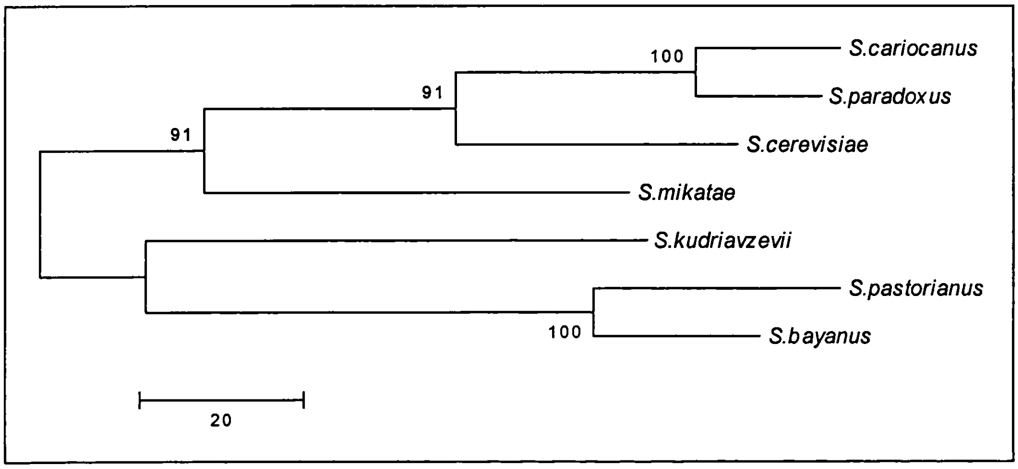

Figure 5.3: Maximum parsimony tree of *FAL1* gene sequences. Strains used: *S. cerevisiae* CBS 8803, *S. bayanus* CBS 7001, *S. kudriavzevii* NCYC 2889, *S. mikatae* NCYC 2888, *S. paradoxus* NCYC 2600, *S. pastorianus* W-34/70 and *S. cariocanus* NCYC 2890 (Number of bootstrap reps = 100).

The PCR results from this species are a typical example of those for the *Saccharomyces* group as a whole. Standard results are shown in figures 5.5 and 5.6. Primer pairs F1 / R2 and F2 / R2 were species specific. The non-specific amplification shown in figure 5.6 would be very simple to remove by raising the optimal annealing temperature as revealed by a temperature gradient PCR. This was unnecessary due to two other primer pairs being specific.

```
Smik F1                  ACAAGCAATTGATTTGAGGAAAAG
S.cerevisiae    FAL1     ......T........A.....G.A
S.bayanus       FAL1     ......T.....C....AA....A
S.mikatae       FAL1     ........................
S.paradoxus     FAL1     ......T........A.....G.A
S.kudriavzevii  FAL1     ......T.....CC...AA.G..T
S.cariocanus    FAL1     ......T........A.....G.A

Smik F2                  AGGGATTTGCAAGCTTTGATATTG
S.cerevisiae    FAL1     .AA.....A......C.A...C.A
S.bayanus       FAL1     .A............AC.....C.A
S.mikatae       FAL1     ........................
S.paradoxus     FAL1     .AA............C.A...C.A
S.kudriavzevii  FAL1     .T......A..G...C.A.....A
S.cariocanus    FAL1     .AA..............A...C.A

Smik R1                  CCAGTCTTCTTTGTCAACGTTG
S.cerevisiae    FAL1     ...C..............A..A
S.bayanus       FAL1     ............A.....A...
S.mikatae       FAL1     ......................
S.paradoxus     FAL1     ..................A..A
S.kudriavzevii  FAL1     ...A...........G..A...
S.cariocanus    FAL1     ..................A..A

Smik R2                  CACGAGAATGACCTGTCCTAAAT
S.cerevisiae    FAL1     ................T.....A
S.bayanus       FAL1     ..........T..G..T.....A
S.mikatae       FAL1     .......................
S.paradoxus     FAL1     ....G.....G.....T..G..A
S.kudriavzevii  FAL1     .T...........G..T.....G
S.cariocanus    FAL1     ....G.....G.....T..G..A
```

Figure 5.4: Potentially *S. mikatae* specific primer target loci in each of the *Saccharomyces sensu stricto* species.

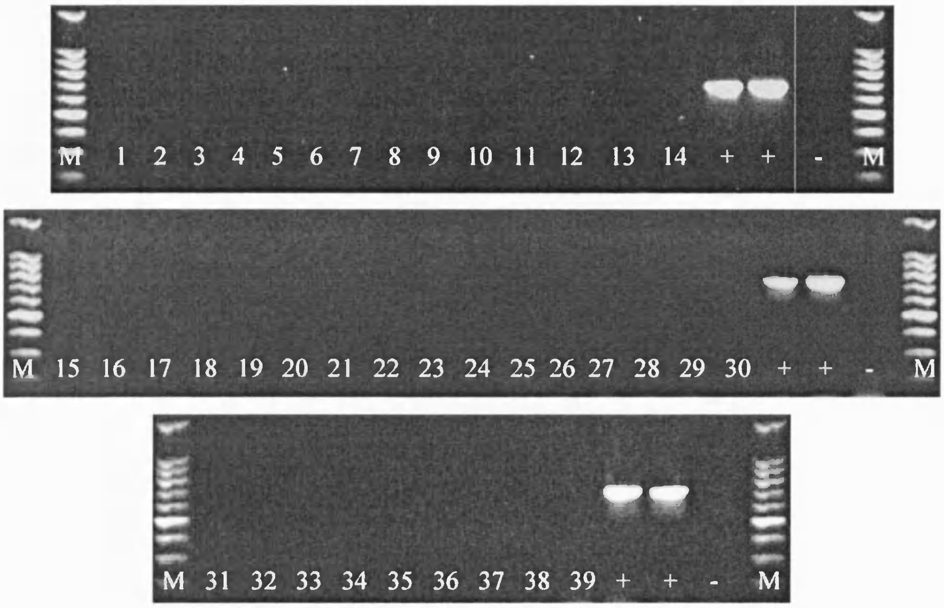

Figure 5.5: Example of a typical specific primer pair result. *S. mikatae* primers F1 / R2 with (1 – 8) *S. cerevisiae* 1 – 8, (9 – 14) *S. bayanus* 1 – 6, (15 – 21) *S. paradoxus* 1 – 6, (22 – 30) *S. cariocanus* 3 – 11, (31 – 35) *S. pastorianus* 1 – 4, (36 – 37) *S. kudriavzevii* 1 & 2, (38 – 39) *S. cariocanus* 1 & 2, (+ & +) *S. mikatae* 1 & 2.

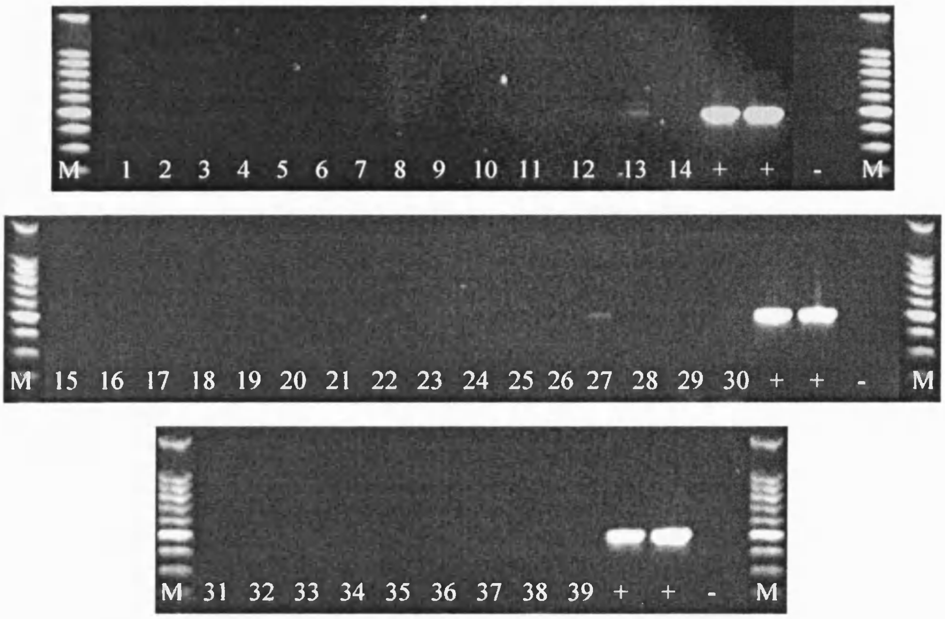

Figure 5.6: Example of a typical non-specific primer pair result. *S. mikatae* primers F2 / R1 with (1 – 8) *S. cerevisiae* 1 – 8, (9 – 14) *S. bayanus* 1 – 6, (15 – 21) *S. paradoxus* 1 – 6, (22 – 30) *S. cariocanus* 3 – 11, (31 – 35) *S. pastorianus* 1 – 4, (36 – 37) *S. kudriavzevii* 1 & 2, (38 – 39) *S. cariocanus* 1 & 2, (+ & +) *S. mikatae* 1 & 2.

```
                       10        20        30        40        50        60        70        80        90       100
               ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.cerevisiae   MSGFHNVGNINMMAQQQMQQNRIKISVRNWQNATMNDLINFISRNARVAVYDAHVEGPLVIGYVNSKAEAESLMKWNGVRFAGSNLKFELLDNNGASAGT
S.bayanus      .................L....V.V.I.......................L....D....V.F.............................G..G....
S.mikatae      ........................V....................V...IL................T........................G.......
S.paradoxus    ........................V.I......................IF.........V...............................G.......
S.pastorianus  .................L....V.V.I.......................L....D....V.F.............................G..G....
S.cariocanus   ------------------------------------------------..F.........V...............................G.......
                      110       120       130       140       150       160       170       180       190       200
               ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.cerevisiae   SDTISFLRGVLLKRYDPQTKLLNLGALHSDPELIQKGVFSSISTQSKMFPAMMKLASTEKSLIVESVNLADNQLKDISAISTLAQTFPNLKNLCLANNQI
S.bayanus      ..................S........................................R..V......................S..............
S.mikatae      ...........................................................R........................................
S.paradoxus    ...........................................................R..V.....................................
S.pastorianus  ..................S...X..........--------------............P..V......................S..............
S.cariocanus   ...........................................................R..V.....................................
                      410       420       430       440       450       460       470       480       490       500
               ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.cerevisiae   GFFEETGKPELESNKKTGKNNYQKNRRYNHGYNSTSNNKLSKKSFDRTWVIVPMNNSVIIASDLLTVRAYSTGAWKTASIAIXXQPPQQQASVLPQVASM
S.bayanus      ..........M.A...........................F..............................A.......DGVPASL.L..P.....A.IP
S.mikatae      ............P................................................................T..EV....VP..P.A.......
S.paradoxus    .....I......P.............................................................................S.........
S.pastorianus  ..........M.A---------------------------------------------------------------------------------------
S.cariocanus   ----------------------------------------------------------------------------------------------------
                      510       520       530       540       550       560       570       580       590       600
               ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.cerevisiae   -----NPNITTPPQPQPSVVPGGMSIPGAPQGA-MVMAPTLQLPPDVQSRLNPVQLELLNKLHLETKLNAEYTFMLAEQSNWNYEVAIKGFQSSMNGIPR
S.bayanus      PNASM........PAP........NM..VG..PP..............................Q...........................G.......
S.mikatae      -SASM..T..............A.NM..T...P-..................................................................
S.paradoxus    -----...........S.......NM......P-.I...................................................V......V.....
S.pastorianus  ----M........P..........NM..V...P-..............................Q...........................G.......
S.cariocanus   ------------....S.......NM......P-.I......................------------------------------------------
```

Figure 5.7: ClustalW amino acid alignment of the *Saccharomyces sensu stricto* species partial *MEX67* gene. *S. cerevisiae* tested primer loci are highlighted.

<u>*S. cerevisiae*</u>

S. cerevisiae primers were targeted to the *MEX67* gene. A ClustalW amino acid alignment of the *Saccharomyces MEX67* sequences available is shown in figure 5.7. Figure 5.8 shows the sequence relationship of this gene in the *Saccharomyces* genus using a maximum parsimony phylogenetic tree of the genus. *S. cerevisiae* had 5 potentially specific primer pairs: F1 / R1, F1 / R2, F2 / R1, F2 / R2 and F3 / R1 (F3 and R2 were targeted to the same region of the gene on opposite strands and therefore could not be paired). Primer sequences are listed in table 5.3 and their target sites in other species are shown in figure 5.9.

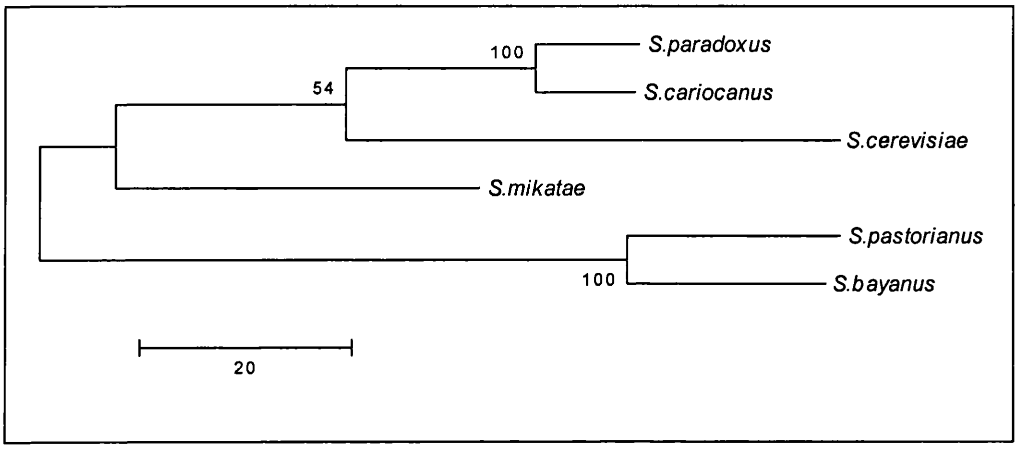

Figure 5.8: Maximum parsimony tree of *MEX67* genes. Strains used: *S. cerevisiae* CBS 8803, *S. bayanus* CBS 7001, *S. mikatae* NCYC 2888, *S. paradoxus* NCYC 2600, *S. pastorianus* W-34/70 and *S. cariocanus* NCYC 2890 (Number of bootstrap reps = 100).

Primer pairs were designed without comparison to the *S. kudriavzevii MEX67* sequence. Amplicons could not be consistently produced and sequenced with degenerate primers. Despite the lack of this sequence, successful specific primers were designed. Only F2 / R2 was specific, the other primer pairs presented typical problems that can occur with any pair (figure 5.10). *S. paradoxus* strains produced an oversized product with F1 / R1 primers.

```
Scer F1                    TAGATAACAATGGGGCCTCTGCG
S.cerevisiae MEX67         .......................
S.bayanus    MEX67         .G...GG......A.GT.....T
S.mikatae    MEX67         ....CGG......A..T.....A
S.paradoxus  MEX67         ....CGG......A........A
S.cariocanus MEX67         ....CGG......A........A

Scer F2                    GCGCTTTACATTCAGATCCCGAG
S.cerevisiae MEX67         .......................
S.bayanus    MEX67         .A..A........G..C..T..A
S.mikatae    MEX67         .A..C........T..C..T..A
S.paradoxus  MEX67         .G..C........T........A
S.cariocanus MEX67         .G..C........T........A

Scer F3                    TCAATCTTGCTGACAACCAACTT
S.cerevisiae MEX67         .......................
S.bayanus    MEX67         .G........G..T........A
S.mikatae    MEX67         .G...........T........A
S.paradoxus  MEX67         .G...........T........A
S.cariocanus MEX67         .G...........T........A

Scer R1                    GTGCGCCGGGAATGCTCATCCCT
S.cerevisiae MEX67         .......................
S.bayanus    MEX67         CCA.......C...T....T..C
S.mikatae    MEX67         .C.TA.....C...T....TG.C
S.paradoxus  MEX67         ..........C...T.......C
S.cariocanus MEX67         .C........C...T.......C

Scer R2                    TAAGTTGGTTGTCAGCAAGATTG
S.cerevisiae MEX67         .......................
S.bayanus    MEX67         .T........A..C........C
S.mikatae    MEX67         .T........A...........C
S.paradoxus  MEX67         .T........A...........C
S.cariocanus MEX67         .T........A...........C
```

Figure 5.9: *S. cerevisiae* potentially specific primer target loci in each of the *Saccharomyces sensu stricto* species.

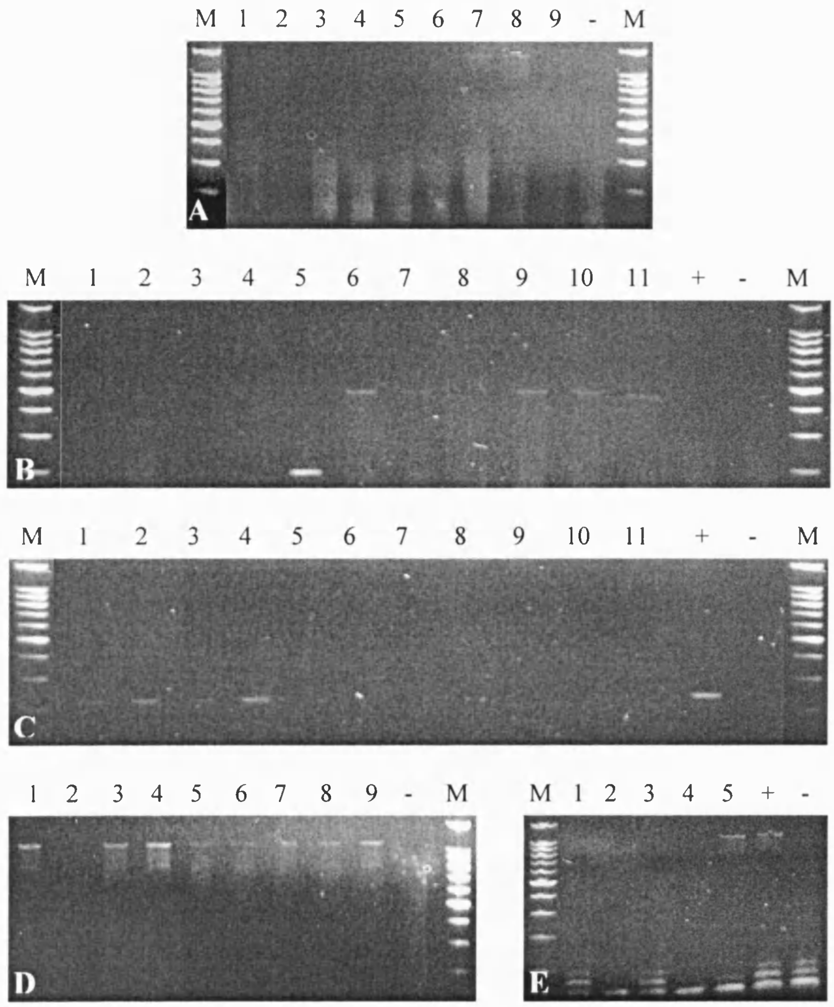

Figure 5.10: Potentially specific *S.cerevisiae* primer pair testing.

(A) F1 / R1 (1 – 9) *S. cerevisiae* 1 – 9,

(B) F1 / R1 (1 – 5) *S.bayanus* 1 – 5 (6 – 11) *S.paradoxus* 1 – 6,

(C) F1 / R2 R1 (1 – 5) *S.bayanus* 1 – 5 (6 – 11) *S.paradoxus* 1 – 6,

(D) F3 / R1 (1 – 9) *S. cerevisiae* 1 – 9,

(E) F3 / R1 (1) *S. pastorianus* 1, (2 – 3) *S. mikatae* 1 – 2, (4 – 5) *S.kudriavzevii* 1–2.

Typical problems represented are: (A) and (D) Target strains not amplified. (B) Non-target species amplified at the MEX67 target site and additional sites. (C) Non-target amplification at the MEX67 site. (E) Multiple non-target products amplified, likely to be primer artifacts as they occur in the negative control.

```
                         10        20        30        40        50        60        70        80        90       100
                 ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.kudriavzevii   MADSVKVSGPSTSDMEYYYKSLYPFKHVFNWLNHSPKPSRDMINREFAMAFRSGAYKRYNSFNSVQEFKAQIERSNPDRFEIGAIYNKPPRERDTLLKSE
S.cerevisiae     .TN...TN...S...............I......................................D......KA.........................
S.bayanus        .IE...AN...S...............I.............................................KG.........................
S.mikatae        .T..I.AN...S...............I..............................................A.........................
S.paradoxus      .T....AN...S...............I.............................................KA.........................
S.pastorianus    .VE...AN...S...............I.............................................KG.........................
S.cariocanus     ------------------------------------------------------...................KA.........V...............
                        110       120       130       140       150       160       170       180       190
200
                 ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.kudriavzevii   LKALEKELVFDIDMDDYDSFRTCCSGAQVCFKCWKFISLAMKIMNTALTEDLGYKDLIWVFSGRRGAHCWVSDKRARALTDIQRRNVLDYVNVVRDRNTD
S.cerevisiae     ..................A...........S............T....R..F....F........................V...........I......
S.bayanus        ..P...............A................................F...NF..I........................................
S.mikatae        ..................A...........S....................F....F........................L................S.
S.paradoxus      ..................A...........C....................F....F........................V..................
S.pastorianus    .RP...............A................................F...NF..I........................................
S.cariocanus     ..................A...........C....................F....F........................V..................
                        210       220       230       240       250       260       270       280       290       300
                 ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.kudriavzevii   KRLALKRPYHPHLARSLEQLKPFFISIMLEEQNPWEDDQHAIDTLLPALYDKQLIDSLRKHWLDNPGRSSKEKWNDIDRVATSLFK--------------
S.cerevisiae     ........................V.................Q...............K.Y.....R...........QI......GPKQDSHIIKLREC
S.bayanus        ........................V..........Q......Q............E....Y.................Q..A....GPKQDSHVARLREC
S.mikatae        R.........................................Q......H........K.Y..................L.S....GLKQDSHISRLREC
S.paradoxus      ........................V.................Q...............K.Y.....R...........QI......GPKQDSHISRLREC
S.pastorianus    .................D......V..........Q......Q............E....Y..E..............Q.......GPKQDSHIARLREC
S.cariocanus     ........................V.................Q...............K.Y..E..............QI......GPKQDSHISRLREC
                        210       220       230       240       250       260       270       280       290       300
                 ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.kudriavzevii   KRLALKRPYHPHLARSLEQLKPFFISIMLEEQNPWEDDQHAIDTLLPALYDKQLIDSLRKHWLDNPGRSSKEKWNDIDRVATSLFKGP----HIGRLREC
S.cerevisiae     ........................V.................Q...............K.Y.....R...........QI........KQDS..IK....
S.bayanus        ........................V..........Q......Q............E....Y.................Q..A......KQDS.VA.....
S.mikatae        R.........................................Q......H........K.Y..................L.S.....LKQDS..S.....
S.paradoxus      ........................V.................Q...............K.Y.....R...........QI........KQDS..S.....
S.pastorianus    .................D......V..........Q......Q............E....Y..E..............Q.........KQDS..A.....
S.cariocanus     ........................V.................Q...............K.Y..E..............QI........KQDS..S.....
```

Figure 5.11: ClustalW amino acid alignment of the *Saccharomyces sensu stricto* species partial *PRI1* gene. *S kudriavzevii* tested primer loci are highlighted.

S. kudriavzevii

S. kudriavzevii primers were targeted to the *PRI1* gene. A ClustalW amino acid alignment of this gene is shown in figure 5.11. Figure 5.12 illustrates the sequence relationships of this gene among *Saccharomyces* species via a maximum parsimony phylogenetic tree.

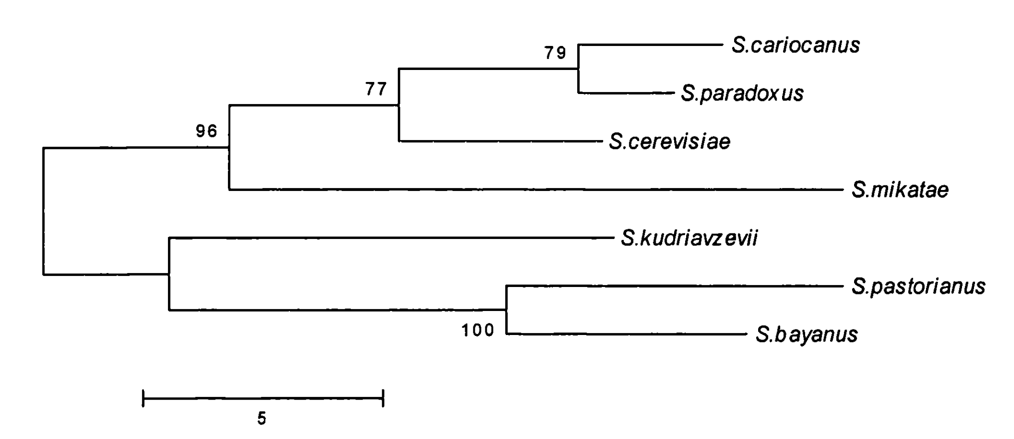

Figure 5.12: Maximum parsimony tree of *PRI1* genes. Strains used: *S. cerevisiae* CBS 8803, *S. bayanus* CBS 7001, *S. kudriavzevii* NCYC 2889, *S. mikatae* NCYC 2888, *S. paradoxus* NCYC 2600, *S. pastorianus* W-34/70 and *S. cariocanus* NCYC 2890 (Number of bootstrap reps = 100).

Three forward and two reverse primers were designed, their sequences are listed in table 5.3 and target sites are shown in figure 5.13. The R1 primer was designed to encompass a target species-specific gap in the alignment.

Similarly to *S. cerevisiae*, the *S. kudriavzevii* experiments produced only one specific primer pair: F2 / R1. Table 5.5 shows the species with which each primer pair was positive or negative.

	Test							
Primer	Sample	Target	Sibling	*S. cariocanus*	*S. cerevisiae*	*S. mikatae*	*S. paradoxus*	*S. pastorianus*
F1 / R1	✓	✓	-	-	-	-	-	-
F2 / R1	✗	✓	✗	✗	✗	✗	✗	✗
F3 / R1	✓	✗	-	-	-	-	-	-
F1 / R2	✓	✓	-	-	-	-	-	-
F2 / R2	✓	✓	-	-	-	-	-	-
F3 / R2	✓	✓	-	-	-	-	-	-

Table 5.5: The results with *S. kudriavzevii* primer pairs. The sibling is *S. bayanus*. Each pair was first tested on a sample of isolates and the target isolates. Only F2/R1 was provisionally specific and was therefore the only pair tested on all other *Saccharomyces* isolates. F2/R1 is the *S. kudriavzevii* specific primer pair. (✓) Amplification, (✗) no amplification, (-) the experiment was not performed.

```
Skud F1                     TTCAAGGCACAGATAGAGAGAAGT
S.cerevisiae    PRI1        .....A.....A.....A.A.GCC
S.bayanus       PRI1        .....A..G.....T..A.A.G.C
S.mikatae       PRI1        .....A..G........A...GCC
S.paradoxus     PRI1        .................A.A.GCC
S.kudriavzevii  PRI1        ........................
S.cariocanus    PRI1        .................A.A.GCC

Skud F2                     ATCTATAACAAACCGCCAAGGGAG
S.cerevisiae    PRI1        ..A.................A..A
S.bayanus       PRI1        ..T..C..T..............A
S.mikatae       PRI1        ..A....................A
S.paradoxus     PRI1        ..A.................A..A
S.kudriavzevii  PRI1        ........................
S.cariocanus    PRI1        G.A...........C.....A..A

Skud F3                     ACGGAAGATCTTGGTTACAAAGAT
S.cerevisiae    PRI1        .GA..G...T.......T..G..C
S.bayanus       PRI1        .........T.C........GA.C
S.mikatae       PRI1        ..A......T....C..T..G..C
S.paradoxus     PRI1        ..A..G...T.......T..G..C
S.kudriavzevii  PRI1        ........................
S.cariocanus    PRI1        ..A..G...T.......T..G..C

Skud R1                     CGTAACCTACCTATATGAGGG------------CCT
S.cerevisiae    PRI1        ......T..AT...G....A.TCTTGCTTGGGG...
S.bayanus       PRI1        ..C.....GG...C.....A.TCTTGCTTGGGA...
S.mikatae       PRI1        ........G.T...G....A.TCTTGCTTGAGA...
S.paradoxus     PRI1        ....G...G.T......G.A.TCTTGCTTGGGA..C
S.kudriavzevii  PRI1        .....................------------...
S.cariocanus    PRI1        ....G...G.T......G.A.TCTTGCTTGGGA...

Skud R2                     TCTAACCAATGCTTTCTCAATGAG
S.cerevisiae    PRI1        ..C.G.....A...CT.T.....A
S.bayanus       PRI1        ..C.......A......T.....T
S.mikatae       PRI1        ..C.......AT..CT.T.....A
S.paradoxus     PRI1        ..C.G.....A...CT.T.....A
S.kudriavzevii  PRI1        ........................
S.cariocanus    PRI1        ..C.G.....A...CT.T.....A
```

Figure 5.13: *S. kudriavzevii* potentially specific primer target loci in each of the *Saccharomyces sensu stricto* species.

```
                          10        20        30        40        50        60        70        80        90       100
                 ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.cariocanus     KTKDMETFYPFSMLETGWDILFFWVTRMILLGLKLTGSIPFKEVFCHSLVRDAQGRKMSKSLGNVIDPLDVISGIKLDDLYAKLLQGNLDPREVEKAKIG
S.cerevisiae     ......................................V.................................T.......H...................
S.bayanus        ......................................V..........................V......T.......H...................
S.mikatae        ......................................V.................................A....N..HD...H..............
S.paradoxus      ......................................V.............................................................
S.kudriavzevii   ......................................V..........................V......A.......H...S........I....A.
S.pastorianus    ..M......................S.......................................V......T....E..H...................
            110       120       130       140       150       160       170       180       190       200
                 ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.cariocanus     QKESYPNGIPQCGTDAMRFALCAYTTGGRDINLDILRVEGYRKFCNKIYQATKFALMRLGDDYQPPTIEGLSGNESLVEKWILHKLTETSKIVNEALDKR
S.cerevisiae     ..................................................................AT................................
S.bayanus        ..................................................................AN................................
S.mikatae        ..................................................................AN................................
S.paradoxus      ...............................................................E..AT................................
S.kudriavzevii   .R...............................................................SAT................................
S.pastorianus    ......................................................G...........AN.......................V......R.
                          210       220       230       240       250       260       270       280       290       300
                 ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.cariocanus     DFLTSTSSIYEFWYLICDVYIENSKYLIQEGSEVEKKSAKDTLYILLDNALKLIHPFMPFISEEMWQRLPKRSTEKAASIVKASYPVYVSQYDDVKSADA
S.cerevisiae     ................................AI........................................................E.......N.
S.bayanus        .......T........................A..Q................................I........VT....A....E.K...A...N.
S.mikatae        ................................A..Q.................................V.....T......E...D.A.N.
S.paradoxus      ................................V.......................................................I.........N.
S.kudriavzevii   .......A........................N..................R................I..............A....K.E...D...N.
S.pastorianus    .......I........................A..Q................................I.........T....A...HE.K.........
                          310       320       330       340       350       360       370       380       390       400
                 ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.cariocanus     YDLVLNITKEARSLLSEYNILKNGKVFVESNHEEYFKTAKDQKDSIVSLIKAIDEVTVVRDASEIPEGCVLQSVNPEVNVHPSX----------------
S.cerevisiae     .......................................E.........................................LLVKGHVDIDAEIAKVQKK
S.bayanus        ...I...........................L....G..NA.................HES.........K.........LLVKGHVDIDAEITKVQKK
S.mikatae        ...............................RD......N.....................V...................LLVKGHVDIDAEIAKVQKK
S.paradoxus      .................................................................................LLVKGHVDIDAEIAKVQKK
S.kudriavzevii   ...I...............................Y...N................ALA............K.........LLVKGHVDIDAEISKVQKK
S.pastorianus    .E.............................LK...E..NA.................HES.........K.........LLVKGHVDIDAEITKVQKK
```

Figure 5.14: ClustalW amino acid alignment of the *Saccharomyces sensu stricto* species partial *VAS1* gene. *S. cariocanus* tested primer loci are highlighted.

S. cariocanus

S. cariocanus primers were targeted to the *VAS1* gene. A ClustalW amino acid alignment for the 3' end of this gene is shown in figure 5.14. Figure 5.15 illustrates the sequence relationships of this gene among *Saccharomyces* species via a maximum parsimony phylogenetic tree.

The *VAS1* gene is approximately 3.3kb long. Only the 3' end was used to spare the effort of sequencing the full length for *S. cariocanus*. The 1.1kb *S. cariocanus* amplicon was obtained by sequencing both strands of an amplicon produced using degenerate primers.

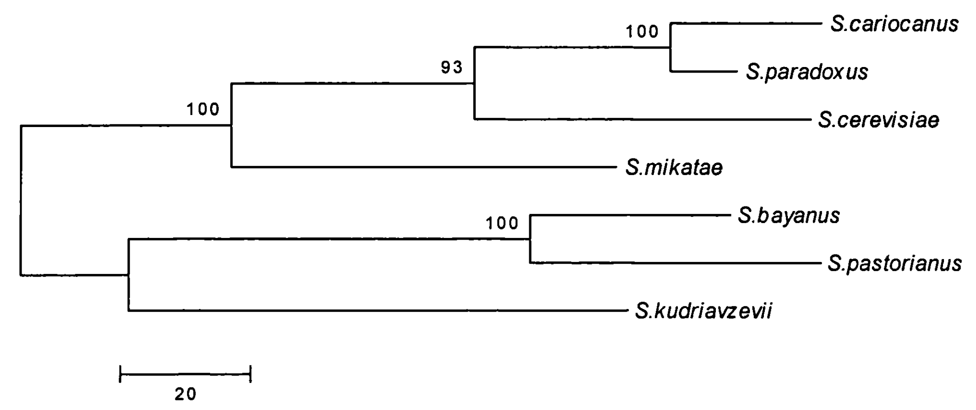

Figure 5.15: Maximum parsimony tree of *VAS1* genes. Strains used: *S. cerevisiae* CBS 8803, *S. bayanus* CBS 7001, *S. kudriavzevii* NCYC 2889, *S. mikatae* NCYC 2888, *S. paradoxus* NCYC 2600, *S. pastorianus* W-34/70 and *S. cariocanus* NCYC 2890 (Number of bootstrap reps = 100).

Three forward primers and three reverse primers were designed, allowing for nine combinations of potentially specific primer pairs. Their sequences are listed in table 5.3 and target sites are shown in figure 5.16. Table 5.6 shows the species with which each primer pair was positive or negative.

	Test							
Primer	Sample	Target	Sibling	*S. kudriavzevii*	*S. cerevisiae*	*S. mikatae*	*S. bayanus*	*S. pastorianus*
F1 / R1	✗	✓	✗	✗	✗	✗	✗	✗
F2 / R1	✓	✓	-	-	-	-	-	-
F3 / R1	✗	✓	✗	✗	✗	✗	✗	✗
F1 / R2	✗	✓	✗	✗	✗	✗	✗	✓
F2 / R2	✗	✓	✗	✗	✗	✗	✗	✗
F3 / R2	✓	✓	-	-	-	-	-	-
F1 / R3	✗	✓	✗	✗	✗	✗	✗	✗
F2 / R3	✓	✓	-	-	-	-	-	-
F3 / R3	✗	✓	✗	✗	✓	✗	✓	✓

Table 5.6: The results with *S. cariocanus* primer pairs. The sibling is *S. paradoxus*. Each pair was first tested on a sample of isolates and the target isolates. F1 / R1, F3 / R1, F2 / R2 and F1 / R3 are *S. cariocanus* specific. (✓) Amplification, (✗) no amplification, (-) the experiment was not performed.

```
Scar F1                              GGGCTTAAAATTAACTGGTTCAA
S.cerevisiae      VAS1               ............G..C......G
S.bayanus         VAS1               ......G...........C..CG
S.mikatae         VAS1               ...T..G.....G........TG
S.paradoxus       VAS1               ............G..C......G
S.kudriavzevii    VAS1               ...T.....G..G........CG
S.cariocanus      VAS1               .......................

Scar F2                              GACCATTATCAACCACCTACTAT
S.cerevisiae      VAS1               ...G..............G.C.C
S.bayanus         VAS1               ...G.C.....G......G...A
S.mikatae         VAS1               ..TG....C.........G...A
S.paradoxus       VAS1               ...G.....G........G...C
S.kudriavzevii    VAS1               ..TG...........T..G.C.C
S.cariocanus      VAS1               ...G...................

Scar F3                              GTGATGTCTATATCGAGAACTCT
S.cerevisiae      VAS1               .......T..C............
S.bayanus         VAS1               ..........C.....A..T...
S.mikatae         VAS1               .......T.....T..A.....C
S.paradoxus       VAS1               ...................T..C
S.kudriavzevii    VAS1               ....C.....C..T..A.....C
S.cariocanus      VAS1               .......................

Scar R1                              CACAACCTTCAGGAATTTCAGAG
S.cerevisiae      VAS1               .G........T........G..A
S.bayanus         VAS1               ....T..C..T........G..A
S.mikatae         VAS1               .G..G..C..T...........A
S.paradoxus       VAS1               ..........T........G..A
S.kudriavzevii    VAS1               ....C.....T............
S.cariocanus      VAS1               .......................

Scar R2                              ATCTTTCTGATCCTTAGCGGTT
S.cerevisiae      VAS1               ............T.C...A...
S.bayanus         VAS1               G..C..T..GG.A..G..A...
S.mikatae         VAS1               ......T.....A..G..A...
S.paradoxus       VAS1               ..................A..C
S.kudriavzevii    VAS1               ...C........A.....A...
S.cariocanus      VAS1               ......................

Scar R3                              AGAGGCCTTTACAATTGCGGCG
S.cerevisiae      VAS1               ...A..T..........A...A
S.bayanus         VAS1               ..CA.....G.....G.T.A.A
S.mikatae         VAS1               ..T...T..A.......A.A..
S.paradoxus       VAS1               ......T........C.A...A
S.kudriavzevii    VAS1               ..C......G.....G.A...A
S.cariocanus      VAS1               .................A....
```

Figure 5.16: *S. cariocanus* potentially specific primer target loci in each of the *Saccharomyces sensu stricto* species.

```
                         710       720       730       740       750       760       770       780       790       800
                  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.paradoxus       SKFLEFINLTTSTGNFELATEFLNSLPSDNEEVKTEKARVLIASGKSLPAQNPVTATTTSKAKYTNSKANKNVPVLPTPGMPSTTSIPSVQAPFYGMTPG
S.cerevisiae      A....................................................A..XX........A.T....................M..........
S.bayanus         ........................T........................A.H.-A..A.R.G..S..---....A............SGM.T.....ASV
S.mikatae         .....................................S...V......SP................A.T..K................G.H.....T...
S.kudriavzevii    ........................T......................FSPE.AA.T.A.R....S.A.T.....A..............M.V........
S.pastorianus     ........................T................T.......AV..-A..A.R.G..A.P.V......P..S........SGM.T.Y.....A
S.cariocanus      ....................Q..D.........................V.......A........A.T....................M..........
                         810       820       830       840       850       860       870       880       890       900
                  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.paradoxus       APSGALPPKPYVPATTTTSAPIHTEGKYAPPHQPSMMSPFXXKTSSSTRLNSFAPPPNPYATAAV-PVATASTTSIPQNSFAPIQPGMPIMGDYNAQSNS
S.cerevisiae      .SAN............XX...V.........S....A...VN..N..................T.-.ATNV........T..................S.
S.bayanus         PGT..--......TG.XXAV.V...............S..AN..N.................P.AVST..SA.STF...............N.F....S.
S.mikatae         .ST.....NS.A..A.XXG..V......XX.S.....T..AN..N.F................T-SA...........SA...........S.F..P.S.
S.kudriavzevii    TA.........A.TAAXXG.LAYPD...........TP..AN..N..................A--PI.T.A.L.....A...........N......S.
S.pastorianus     TGA..--......TV.XXAV.V..D...........A...ANR.N..................AASTT.T.APTF................N.F....S.
S.cariocanus      ...............A..............S.........AN....................--------------------------------------
                         910       920       930       940       950       960       970       980       990      1000
                  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.paradoxus       TPSQPPINAVSGQTPHLNRKANDGWNDLPLKVKEKPSRAKAVSVAPPNILSTPTPLNGIHAHATSTMPPPPLSRAPSSVSMVSPPPL-HKTSRVPSLVAT
S.cerevisiae      I..........................................................P.N.A.......................-..N.........
S.bayanus         I.P.AS.S..........................R................AA....VSTY.GNN.....................PP............
S.mikatae         I......S.I...............................................SVP...AN......................-............
S.kudriavzevii    I...L.MS.................---------------------------------------------------------------------------
S.pastorianus     I.PHS.............................R...............AA.SS..VPTY..N...........L..........--..S.........
S.cariocanus      ----------------------------------------------------------------------------------------------------
                        1010      1020      1030      1040      1050      1060      1070      1080      1090      1100
                  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.paradoxus       SESPRASISNPYAPPQPSQQFPVGAPSTTSQTSNAAQATSSNPYAPPPQPRVATPLSGGVPPAPLPKASNPYAPAATTQPNGSSYPPTSAYTNNH--MTS
S.cerevisiae      ................S.....I.TI..AN....T..VA..........Q........................T.............GP.....--...
S.bayanus         .......V......S.SI.-P...T..LK.P.LTPI..P........L........AA...ST..S........T.AA........SNDTRITTPWL.A.
S.mikatae         ..............S.SA..PLM.TS.I......---QAA................L.AI.....S........T.A........A..N..A.T.--V..
S.kudriavzevii    ----------------------------------------------------------------------------------------------------
S.pastorianus     .....V.V......S.ST..P...I..LK.P.PTPI.PP........L........TS...TS......S......AA........S.N.----------
S.cariocanus      -----------------------------------------.....................V..............................S.--.A.
                        1110      1120      1130      1140      1150      1160      1170      1180      1190      1200
                  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.paradoxus       PPPVFNKPPAGPPPISMKKRGNKLASIDQKPSQAATYPPTLSSSASPLQPSQPPTLASQANTSTENVSHEIPADQQPIVDFLKEELARVTPLTPKEYSKQ
S.cerevisiae      .........T..........S......E.N..G..........................V...A....................................
S.bayanus         Q..IS..AA...........S.........QP...AF..P.........T........P........NN.V.............................
S.mikatae         ...IS..L............S..S.N.......ES..............T.........S........R.V............G................
S.kudriavzevii    ---------------------------------------..........A......................P.............S..A..........
S.pastorianus     ----------------------------------------------------------------------------------------------------
S.cariocanus      ...I................S...............................................................................
```

Figure 5.17: ClustalW amino acid alignment of the *Saccharomyces sensu stricto* species partial *SEC31* gene. *S. paradoxus* tested primer loci are highlighted.

S. paradoxus

S. paradoxus potentially specific primers were designed to the *SEC31* gene. A ClustalW amino acid alignment is shown in figure 5.17. Figure 5.18 illustrates the sequence relationships of this gene among *Saccharomyces* species via a maximum parsimony phylogenetic tree. As with *VAS1*, only the 3' end was sequenced because *SEC31* is a long gene at 3.8 kb.

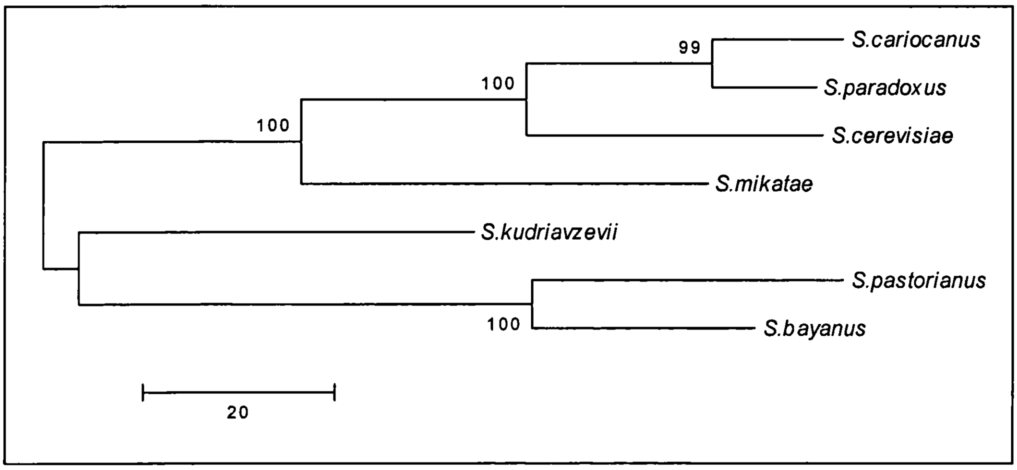

Figure 5.18: Maximum parsimony tree of *SEC31* genes. Strains used: *S. cerevisiae* CBS 8803, *S. bayanus* CBS 7001, *S. kudriavzevii* NCYC 2889, *S. mikatae* NCYC 2888, *S. paradoxus* NCYC 2600, *S. pastorianus* W-34/70 and *S. cariocanus* NCYC 2890 (Number of bootstrap reps = 100).

Two forward and two reverse primers were designed, allowing for four combinations of potentially specific primer pairs. The sequences are listed in table 5.3 and target sites of each primer are shown in figure 5.19. The primer pairs were first tested on all target *S. paradoxus* isolates. The isolates are listed A to P in table 5.7. The strains with culture collection codes beginning "YPS", originated from the study described in which multiple wild strains of *S. cerevisiae* and *S. paradoxus* were isolated from a natural woodland site in Pennsylvania, USA [158]. Matings to genetically marked tester strains were used to establish species identity. When the *S. paradoxus* strains were tested with *S. paradoxus SEC31* potentially specific primer pairs, they were not all amplified. Figure 5.20 [1] shows a result typical of all primer pairs: Strains C, H, K, M, N and O were strongly amplified. Strains A, B, D, E, F, G, I, J, L and P were not amplified or were weakly amplified.

```
Spar F1                         AAGCTAAGTATACAAACTCTAAGG
S.cerevisiae     SEC31          ....C............G.C...A
S.bayanus        SEC31          .G.GC..A..CT....T..C...-
S.mikatae        SEC31          .G...............G.C...A
S.paradoxus      SEC31          ........................
S.kudriavzevii   SEC31          .G.........T....TG.....A
S.cariocanus     SEC31          .................G.G...A

Spar F2                         CTTCTACCACATCTATCCCTAGTG
S.cerevisiae     SEC31          .......T..T.....T......A
S.bayanus        SEC31          .A..C..T..G.....TT..G..A
S.mikatae        SEC31          ....C..T..G.........G...
S.paradoxus      SEC31          ........................
S.kudriavzevii   SEC31          .C..C..T..G..C.........A
S.cariocanus     SEC31          ..........G............A

Spar R1                         GGGCCAGCGGGAGGCTTGTTG
S.cerevisiae     SEC31          .......T......T.....A
S.bayanus        SEC31          ........T.CT.C...A..A
S.mikatae        SEC31          ..T.....T...A....A...
S.paradoxus      SEC31          .....................
S.kudriavzevii   SEC31          ---------------------
S.cariocanus     SEC31          ....................A

Spar R2                         AACCTGGTCGAGCTACTTGTT---TTG
S.cerevisiae     SEC31          ..T............T.....TTG..A
S.bayanus        SEC31          ..T..C..A..A...T.G...TTG..A
S.mikatae        SEC31          ..T.....GA.A...T.....TTG...
S.paradoxus      SEC31          .....................---...
S.kudriavzevii   SEC31          ..T.....A......T.....TTG...
S.cariocanus     SEC31          .....................TTG...
```

Figure 5.19: *S. paradoxus* potentially specific primer target loci in each of the *Saccharomyces sensu stricto* species.

"*S. paradoxus*" isolate	Number	Local Code	Culture collection codes	New identity
A	3	5418	YPS 145	*S. cariocanus*
B	4	5413	YPS 152	*S. cariocanus*
C	1	5103		*S. paradoxus*
D	5	5416	YPS 158	*S. cariocanus*
E	6	2666	YPS 125	*S. cariocanus*
F	7	4259	YPS 138	*S. cariocanus*
G	8	5411	YPS 150	*S. cariocanus*
H	2	4240	CBS 5829	*S. paradoxus*
I	9	5415	YPS 155	*S. cariocanus*
J	10	4241	DBVPG 6304	*S. cariocanus*
K	3	5102		*S. paradoxus*
L		5414		*S. cerevisiae*
M	4	4239	NCYC 2600 CBS 432	*S. paradoxus* (T)
N	5	5101		*S. paradoxus*
O	6	S.bay 4234	CBS 1515	*S. paradoxus*
P	11	5412	YPS 151	*S. cariocanus*

Table 5.7: List of original species assignation of "*S. paradoxus*" isolates that were tested with *S. paradoxus* potentially specific primers. Also listed are the true identification of each isolate as determined by D1 / D2 rDNA sequencing, the strain numbers as listed in table 2.1, local code numbers and culture collection codes. YPS species were obtained from Sniegowski *et al* (2002) [158].

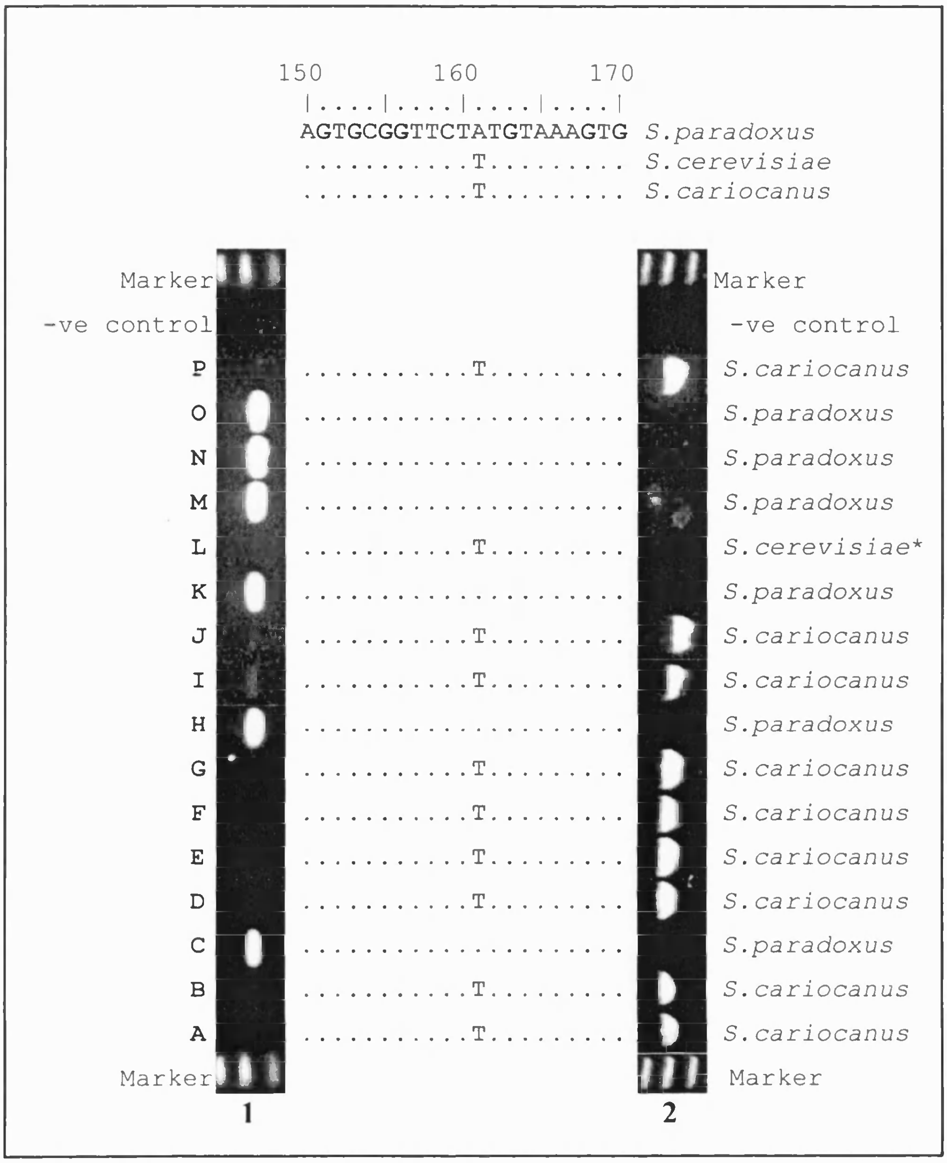

Figure 5.20: Partial D1/D2 sequence alignment of various putative *S.paradoxus* strains illustrating the single SNP between *S.paradoxus* and *S.cariocanus*.
(1) Strains amplified with *S.paradoxus* primers F1/R2
(2) Strains amplified with *S.cariocanus* specific primers F1/R3.
**S.cerevisiae* strain identified by further sequence variation not shown here (see figure 5.21).

The first assumption was that the PCR conditions were too stringent to amplify all of the target strains. This prompted another experiment with a lower annealing temperature of 50°C. The positive and negative results from each experiment were exactly the same as before, although previously weak amplification with the YPS strains increased slightly in strength. When the D1 / D2 sequences were determined,

```
                                     110       120       130       140       150       160       170       180       190       200
                               ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.paradoxus_NCYC_2600          GGACGTCATAGAGGGTGAGAATCCCGTGTGGCGAGGAGTGCGGTTCTATGTAAAGTGCCTTCGAAGAGTCGAGTTGTTTGGGAATGCAGCTCTAAGTGGG
S.cariocanus_NRRL_Y-27337      ...............................................T....................................................
S.cerevisiae_NCYC_505          ...............................................T....................................................
S.paradoxus_A                  ...............................................T....................................................
S.paradoxus_B                  ...............................................T....................................................
S.paradoxus_C                  ....................................................................................................
S.paradoxus_D                  ...............................................T....................................................
S.paradoxus_E                  ...............................................T....................................................
S.paradoxus_F                  ...............................................T....................................................
S.paradoxus_G                  ...............................................T....................................................
S.paradoxus_H                  ....................................................................................................
S.paradoxus_I                  ...............................................T....................................................
S.paradoxus_J                  ...............................................T....................................................
S.paradoxus_K                  ....................................................................................................
S.paradoxus_L                  ...............................................T....................................................
S.paradoxus_M                  ....................................................................................................
S.paradoxus_N                  ....................................................................................................
S.paradoxus_O                  ....................................................................................................
S.paradoxus_P                  ...............................................T....................................................
                                     410       420       430       440       450       460       470       480       490       500
                               ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.paradoxus_NCYC_2600          TCAGTTTTGGTGGCAGGATAAATCCGTAGGAATGTAACTTGCTTCGGGAAGTATTATAGCCTGCGGGAATACTGCCAGCTGGGACTGAGGACTGCGAC
S.cariocanus_NRRL_Y-27337      ..................................................................................................
S.cerevisiae_NCYC_505          .........................A..........G.....C....T...............T..................................
S.paradoxus_A                  ..................................................................................................
S.paradoxus_B                  ..................................................................................................
S.paradoxus_C                  ..................................................................................................
S.paradoxus_D                  ..................................................................................................
S.paradoxus_E                  ..................................................................................................
S.paradoxus_F                  ..................................................................................................
S.paradoxus_G                  ..................................................................................................
S.paradoxus_H                  ..................................................................................................
S.paradoxus_I                  ..................................................................................................
S.paradoxus_J                  ..................................................................................................
S.paradoxus_K                  ..................................................................................................
S.paradoxus_L                  .........................A..........G.....C....T...............T..................................
S.paradoxus_M                  ..................................................................................................
S.paradoxus_N                  ..................................................................................................
S.paradoxus_O                  ..................................................................................................
S.paradoxus_P                  ..................................................................................................
```

Figure 5.21: Two portions of a ClustalW alignment of the D1 / D2 regions of the original "*S. paradoxus*" strains and the type strains of *S. paradoxus*, *S. cariocanus* and *S. cerevisiae*. Only the portions helping with species differentiation are shown.

corresponding to the identity of the *S. cariocanus* type strain (figure 5.21). When the isolates were tested with an *S. cariocanus* specific primer pair, the results were directly opposite to those with *S. paradoxus* primers (figure 5.20 [2]).

The *S. paradoxus* primer pairs were then tested on the other *Saccharomyces* species. The results are listed in table 5.8. Primer pair F1 / R2 did not amplify anything other than target and sibling species whilst all other primer combinations weakly amplified a variety of species. The primer pair F1 / R2 was optimised using a higher annealing temperature to make it species-specific at 65°C (figure 5.22).

	Test							
Primer	Sample	Target	Sibling	*S. bayanus*	*S. cerevisiae*	*S. mikatae*	*S. kudriavzevii*	*S. pastorianus*
F1 / R1	×	✓	✓	✓	×	×	✓	✓
F2 / R1	×	✓	✓	✓	×	×	✓	✓
F1 / R2	×	✓	✓	×	×	×	×	×
F2 / R2	×	✓	✓	✓	×	×	✓	✓

Table 5.8: The results with *S. paradoxus* primer pairs. Sibling species is *S. cariocanus.* Each pair was first tested on a sample of isolates and the target isolates and then tested on all other *Saccharomyces* strains. (✓) Amplification, (×) no amplification, (-) the experiment was not performed.

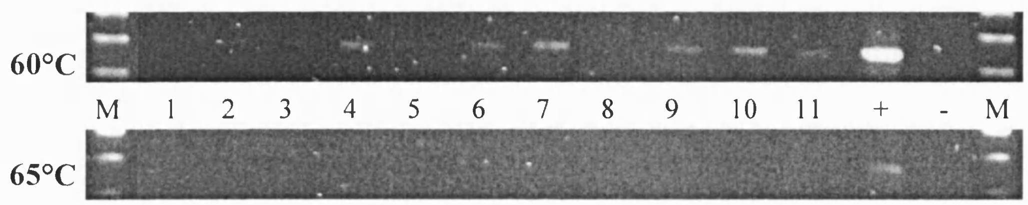

Fig 5.22: *S.cariocanus* strains 1 – 11 amplified with the *S.paradoxus* F1 / R2 primer pair at different annealing temperatures.

```
                          10        20        30        40        50        60        70        80        90       100
                 ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.bayanus        MFASRFDPTQLTAAAAPVTVAPEEPVKVAPEAIVPLKRQASASDSENDDSTGESSDE----GSDXXDDRMQVDYGVSEEDSSEEEEEEEKEEEEKHEEDD
S.cerevisiae     ........S....P.---AS...GI.GTT.P.........TE..N.EYG-SHQD...SSNSS.EEDE........A.......V.-----------..ES
S.mikatae        ........S....PV---AS...KA..PPSRS........TG..D.E..-FRQD...GSNVS.EDD......T..T.......A.-----------.KEN
S.paradoxus      ........S..NTP.---AST...TIRPT.G.........TG..D.D..-FRQD.N.SSS-S.EEDE....IE..AD......A.-----------GKES
S.kudriavzevii   ........S..K..V---.A...KT..PT.Q.........TGPA..D..ASDQE..DSSEKNAG........VS.A...E...V.-----------D.ES
S.pastorianus    ...........SS..-....T......A............T.....S..G.D.....SSGE........................----...KDN...E.
S.cariocanus     -----...C..DTS.---AST...TIRPTSG..........G..D.D..-FLQD.N.SSSTS.EEDE....IE..TD......A.-----------GKES
                         110       120       130       140       150       160       170       180       190       200
                 ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.bayanus        KPSTHNTVLSRFKQTVSLQDKLDASDLVGTKEDESLEENAAPSHQLEQIPQPEFVKNPMNLNRNSSEYKSTGWLNTERVFYDNSMTKPFSDYENQLEPRL
S.cerevisiae     .....S.............ER.G...IAES....GI.DE..ST...K...............T..LQF.........KIY....LI......A.E..AK.
S.mikatae        IS.....................S..IAEK..G....DE.VST.........D............LDF.........KI......IR.......E..A..
S.paradoxus      .....S................NV..IA.NN...R..DYT.ST......................LQF.........KIY...FLI......G.E..AK.
S.kudriavzevii   ....................R.....V..GSG..V..DDTV...K........L...........L.H..A...D...I.............K.E...K.
S.pastorianus    T....T....................I.............T.P..................................................S......
S.cariocanus     ..G..S................NV..IA.NN......DDT.ST...T........R......T-------------------------------------
                         210       220       230       240       250       260       270       280       290       300
                 ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.bayanus        LQNICKNFSTDTFPIQSIILDSVLPILNFTLNISKRNFTKRIGDILVNAATGSGKTLAYSIPVVQTLSKRQINRLRCLIIVPTKLLINQVYTTLAKLTQG
S.cerevisiae     ..........N...........I..V......V......R......................I....F.........I................T.....
S.mikatae        .....N...VN...........I..V........R....R......................I....F.....K.....V..............T.....
S.paradoxus      .........IN...........I..V.....S.......R......................I....F..........................T.....
S.kudriavzevii   .........NA...........I.........V...Y..R....T.......................R.K........V..........H...T.....
S.pastorianus    ..............................................................................................S.....
S.cariocanus     ----------------------------------------------------------------------------------------------------
                         310       320       330       340       350       360       370       380       390       400
                 ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.bayanus        TSLIVSIAKLENSLKDEHSKFLNLEPDILITTPGRLVDHLNMKSINLKNLKFLIIDEADRLLNQSFQGWCPKLMSHLKTDKLDTSPGNVIKMIFSATLTT
S.cerevisiae     ..................K.LS..............................................................L...............
S.mikatae        ..................R.L......................................................Q........F...............
S.paradoxus      .......S..........K.L...............................................................F...............
S.kudriavzevii   M.................K......................................................VF.........F...............
S.pastorianus    ....................................................................................................
S.cariocanus     ----------------------------------------------------------------------------------------------------
                         410       420       430       440       450       460       470       480       490       500
                 ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.bayanus        NTEKLNDLNLYKPKLFLKQTDKLYHLPNKLKEFNINIPTAKSVFKPLILLYSINQFLAHSSTTSKILVFVKSNESSIRLSKLLQLINETRSQSSILTNLR
S.cerevisiae     ......G.................Q.....N............Y.........C..M...PIAA...I..................C.S.....V.K..Q
S.mikatae        ........................Q...........V......Y.........C..I..PP.AA......................C.S....DM.RK.Q
S.paradoxus      ......G.................Q.....D............Y........VC..IT..P.AA...I..................C.S....NM.K..Q
S.kudriavzevii   ........................Q..S..R.....V.....IY.........D..I..L.DA....I...........T......S.S....N.FR..Q
S.pastorianus    ........................................................M...NA.......................M..............
S.cariocanus     ----------------------------------------------------------------------------------------------------
```

Figure 5.23: ClustalW amino acid alignment of *the Saccharomyces sensu stricto* species partial *DBP6* gene. *S. bayanus* tested primer loci are highlighted.

S. bayanus

S. bayanus primers were targeted to the *DBP6* gene. A ClustalW amino acid alignment is shown in figure 5.23. Only the 5' end of this gene was used as it showed a high level of diversity. The 3' end was highly conserved at both the amino acid and nucleic acid level and would have been unsuitable for primer design. Figure 5.24 illustrates the sequence relationships of this gene among *Saccharomyces* species via a maximum parsimony phylogenetic tree of the genus.

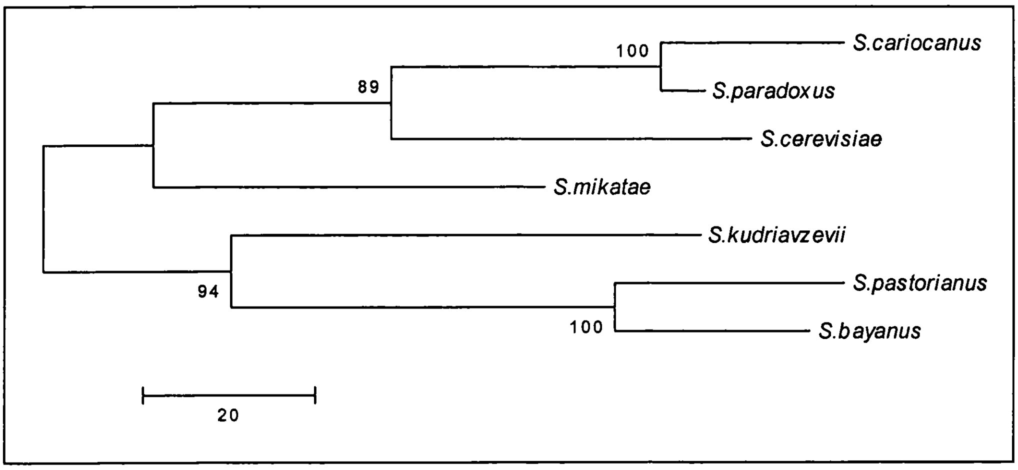

Figure 5.24: Maximum parsimony tree of *DBP6* genes. Strains used: *S. cerevisiae* CBS 8803, *S. bayanus* CBS 7001, *S. kudriavzevii* NCYC 2889, *S. mikatae* NCYC 2888, *S. paradoxus* NCYC 2600, *S. pastorianus* W-34/70 and *S. cariocanus* NCYC 2890 (Number of bootstrap reps = 100).

Two forward and two reverse primers were designed that encompassed gaps in the DBP6 sequence. Primer sequences are listed in table 5.3 and target sites are shown in figure 5.25. However, there were only three combinations of primer pairs as F2 and R2 targeted the same region but on opposite strands. The results from the primer testing experiments are listed in table 5.9.

```
          Sbay F1                         GCTGACTGCTGCTGCTGCCCCCG
          S.cerevisiae    DBP6  AT........C....A-------
          S.bayanus       DBP6  .......................
          S.mikatae       DBP6  A.........C.C.T.-------
          S.paradoxus     DBP6  AT...A.A..C.....-------
          S.kudriavzevii DBP6   .T...AA.......TG-------
          S.cariocanus    DBP6  ....GA.A..T.....-------

Sbay F2                    A-----------AGGCTCCGATA---ATGATGACCG
S.cerevisiae     DBP6   .GAGTTCCAACTCTA.T..A..GGAAG.C..A.....
S.bayanus        DBP6   .------------...........---..........
S.mikatae        DBP6   .AGGCTCTAACGT.A.T..G..GGATG..........
S.paradoxus      DBP6   .GAGCTCCAG---TA.T..A..AGAAG....G.....
S.kudriavzevii DBP6     .TAGTTCCGAGAA.AA.G.A.GA.---..........
S.cariocanus     DBP6   .GAGCTCCAGCACTA.T..A..AGAAG....G.....

          Sbay R1                   TGTTATGAGTACTTGGTTTGTCG
          S.cerevisiae    DBP6  ..C.G...........C..ACTC
          S.bayanus       DBP6  .......................
          S.mikatae       DBP6  ...............A.A...TC
          S.paradoxus     DBP6  ..C....G...........ACTC
          S.kudriavzevii DBP6   .......G..G..C......CTC
          S.cariocanus    DBP6  ..C....G....C......ACTC

Sbay R2                    ACGGTGATCAT---TATCGGAGCCT------------T
S.cerevisiae     DBP6   .....CT..G.CTTCC..T..A.TAGAGTTGGAACTC.
S.bayanus        DBP6   .....C.....---...........------------.
S.mikatae        DBP6   .....C.....CATCC..C..A.T.ACGTTAGAGCCT.
S.paradoxus      DBP6   .....CC....CTTCT..T..A.TA---CTGGAGCTC.
S.kudriavzevii DBP6     G....C.....---.TC.T.C.TT.TTCTCGGAACTA.
S.cariocanus     DBP6   .....CC....CTTCT..T..A.TAGTGCTGGAGCTC.
```

Figure 5.25: *S. bayanus* potentially specific primer target sites in each of the *Saccharomyces sensu stricto* species.

	Test							
Primer	Sample	Target	Sibling	*S. paradoxus*	*S. cerevisiae*	*S. mikatae*	*S. cariocanus*	*S. pastorianus*
F1 / R1	×	✓	×	×	✓	×	×	*
F2 / R1	×	✓	×	×	×	×	×	*
F1 / R2	✓	✓	-	-	-	-	-	-

Table 5.9: The results with *S. bayanus* primer pairs. Sibling is *S. kudriavzevii*. Each pair was first tested on a sample of isolates and the target isolates and then tested on all other *Saccharomyces* strains. (✓) Amplification, (×) no amplification, (-) the experiment was not performed, (*) Only *S. pastorianus* 4667 was amplified.

While the experimental results suggested that none of the primer pairs were *S. bayanus* specific, pair F2 / R1 only amplified a single *S. pastorianus* strain (strain 2) and all target strains. This *S. pastorianus* strain was formerly classified as *S. bayanus var. uvarum* and gene sequencing showed that the D1 / D2 rRNA gene was identical to that of *S. bayanus*. The exact nature of the *S. pastorianus* strains was examined further to determine their true identity and if distinguishing them from the other *Saccharomyces* species using this method would be possible.

```
                         510       520       530       540       550       560       570       580       590       600
                  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.pastorianus     FNEKNAGKEEEEPKDYIEDNDDEDIEGDSDGDEAISNLISKIKGEDGDHKLSSKARMIFNDPIFNNVEADLPVNIINDGMMSSESVGDISKLSKKRKHQE
S.cerevisiae      ...GSLE.K...G..........GV.....D....T.....L..QE......................P.....TV...I............N.....E.
S.bayanus         .....V...................................................................Q..........A...L...........
S.mikatae         .D.VS.E.G...KNNF..XX....V.....D....T.....L..QN......................P...TST....I....A.......H.....E.
S.paradoxus       ...G.LE.K...H.........D.V.....D....T.....L..Q.......N...............P......V...V............N.....E.
S.kudriavzevii    ...V..E.G...N......K..........D....T.....L..Q...................T...P.....T......T..A.......H....YE.

                         610       620       630       640       650       660       670       680       690       700
                  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.pastorianus     IQQQEPEEEADSSDESSSDDSDFEIVANDNVSEDFDSDYDSEEERNQTKKEKHSKDIDIATVEAMTLAHQLALGQKNRHDLVDEGFNRYTFRDTEHLPEW
S.cerevisiae      MH.KQ--D......................A..E..........K.........R......................K.................N..D.
S.bayanus         ..K..-......................................K.......................................................
S.mikatae         V..EQ--.....................................K.........R......................K.................N....
S.paradoxus       .H.KQ--D.T.......................E..........K.........R......................K.................N....
S.kudriavzevii    ..P-R--..T..................................K.........R......................K.................N....

                         710       720       730       740       750       760       770       780       790       800
                  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.pastorianus     FLEDEKEHSKINKPITKEAAMAIKEKIKAMNARPIKKVAEAKARKRMRSVARLEKIKKKAGLINDDSDKTEKDKADEISRLMRKVTKKPQAKPKVTLVVA
S.cerevisiae      ................................................A..........................E.............KT.........
S.bayanus         ....................................................................................................
S.mikatae         ................................................A..........................E..............TR........
S.paradoxus       .............................................K..A..........................E.............KT.........
S.kudriavzevii    ...........................R....................A.T........................E.............KT.........

                         810       820       830       840
                  ....|....|....|....|....|....|....|....|....
S.pastorianus     SGKNKGLAGRPKGIKGKYKMVDGVMKNEQRALRRIAKKHHKK*-
S.cerevisiae      ..R..........V............................K*
S.bayanus         Y............V............................*-
S.mikatae         ..R..........V............................K*
S.paradoxus       ..R..........V............................K*
S.kudriavzevii    A.R..........V............................K*
```

Figure 5.26: ClustalW amino acid alignment of *the Saccharomyces sensu stricto* species partial *SPB1* gene. *S. pastorianus* tested primer loci are highlighted.

S. pastorianus

The *SPB1* gene was chosen to design specific primers for *S. pastorianus*. A ClustalW amino acid alignment of the *SPB1* gene is shown in figure 5.26. A maximum parsimony phylogenetic tree of the *SPB1* gene for all *Saccharomyces* target species is shown in figure 5.27.

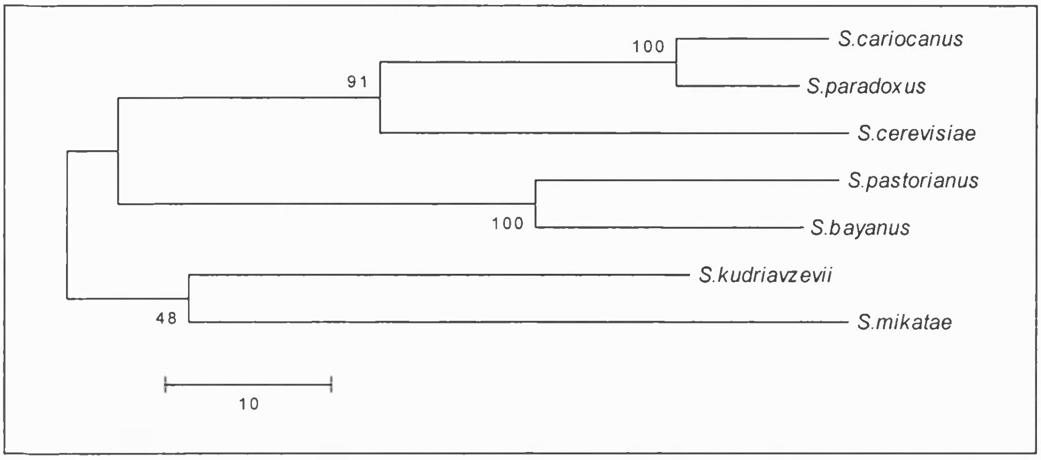

Figure 5.27: Maximum parsimony tree of *SPB1* genes. *S. cerevisiae* CBS 8803, *S. bayanus* CBS 7001, *S. kudriavzevii* NCYC 2889, *S. mikatae* NCYC 2888, *S. paradoxus* NCYC 2600, *S. pastorianus* W-34/70 and *S. cariocanus* NCYC 2890 (Number of bootstrap reps = 100).

Three forward and two reverse primers were designed to be potentially *S. pastorianus* specific. The forward primers encompassed a *S. pastorianus* specific amino acid insertion in the protein sequence. No primers amplified all of the target strains and representative results from testing the primer combinations on the target isolates are shown in figure 5.28.

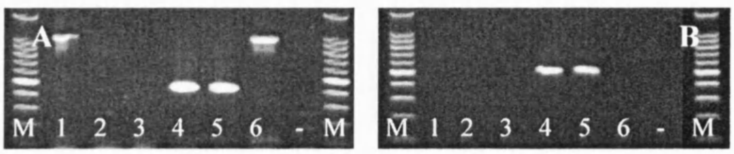

Figure 5.28: Representative results from testing the *S. pastorianus* potentially specific primer pairs on target isolates from the local collection. (A) F2 / R1 [same result as F3 / R1] (B) F2 / R2 [same result as F1 / R1, F1 / R1 and F3 / R2]. (1 – 5) *S. pastorianus* 1 – 5, (6) Local isolate 4246 (since excluded from the study) (-) No DNA.

The results clustered into three subgroups: no product (strains 2 and 3), correct product (strains 4 and 5) and oversized product (strain 1 and local isolate 4246). Each strain (and some additional strains) had their D1 / D2 sequences elucidated. A maximum parsimony bootstrapped tree of the D1 / D2 rDNA gene of *Saccharomyces* species and *S. pastorianus* strains is shown in figure 5.29. The strains form three clusters. A ClustalW alignment of representative strains is shown in figure 5.30. The D1 / D2 sequences of *S. pastorianus* strains 1 and 4 are identical to that of *S. bayanus* type strain CBS 380. The D1 / D2 sequences of *S. pastorianus* strains 3 and 5 are identical to that of *S. cerevisiae*. The D1 / D2 sequence of *S. pastorianus* strain 2 is identical to that of *S. bayanus* CBS 395 (the type strain of *S. bayanus* synonym, *S. uvarum*). More genome sequence analysis was performed in an attempt to accurately place these strains in the *Saccharomyces* phylogeny and to determine if *S. pastorianus* species-specific primers were possible. Degenerate primers were designed to amplify and sequence the *FAL1*, *PRI1* and *MEX67* genes in the *S. pastorianus* isolates. Additional strains in the *S. pastorianus* group were also sequenced [The sequencing was performed by Jonathan Springett, an undergraduate project student]. A list of the strains and their phylogenetic histories is shown in table 5.10.

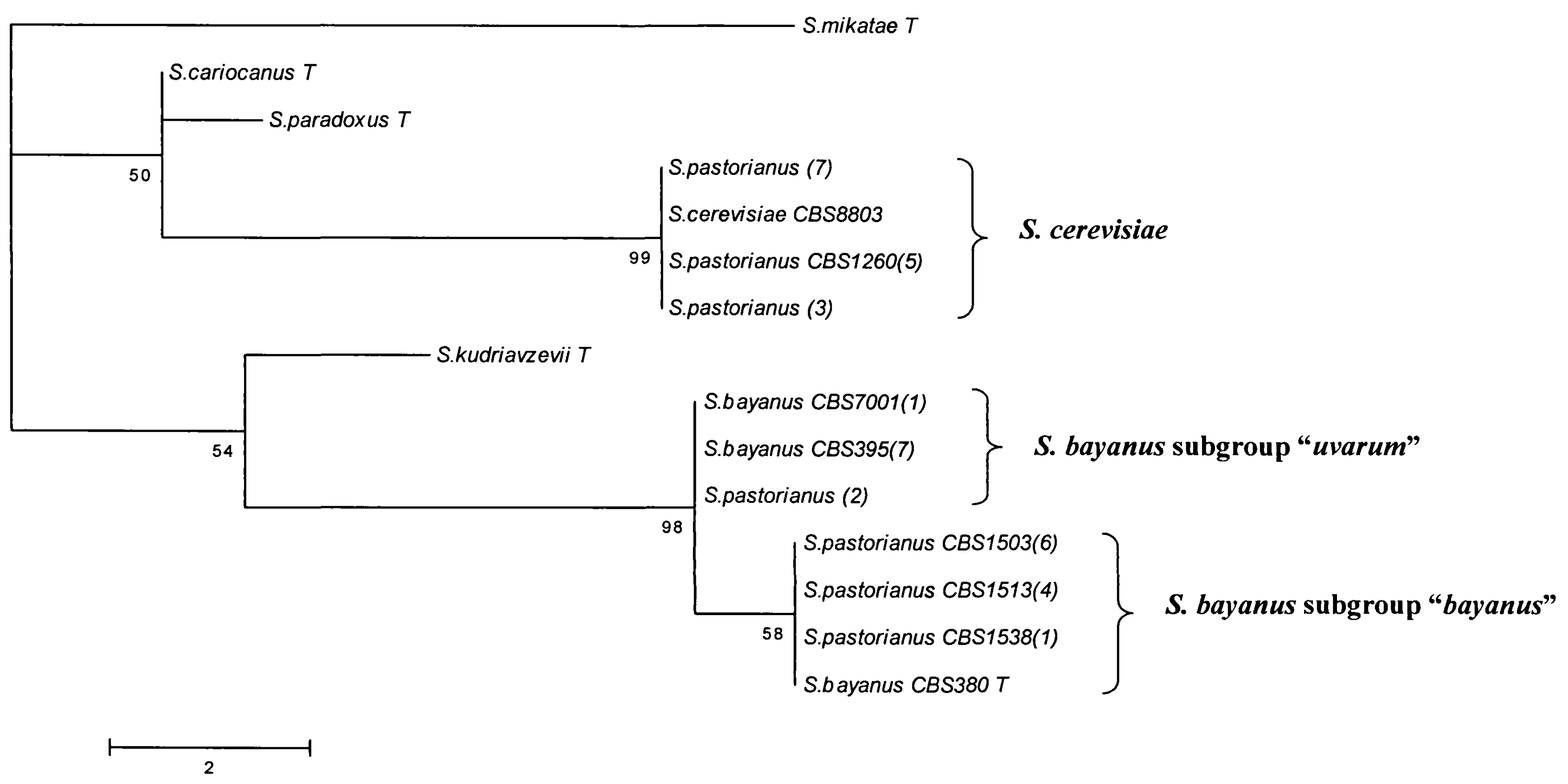

Figure 5.29: Maximum Parsimony phylogenetic tree of the *Saccharomyces* type species, *S. bayanus* strains and *S. pastorianus* strains D1 / D2 rRNA gene. Number of bootstrap replicates = 100. Numbers in parentheses are the strain numbers listed in table 5.10. The strains clustering around *S. bayanus* type strain (CBS 380) exhibit the *S. bayanus* "*bayanus*" subgroup D1 / D2 sequence. The strains that cluster around *S. bayanus* strain 7 (CBS 395) exhibit the D1 / D2 sequence of the more prevalent "*S. uvarum*" subgroup. These subgroups differ in only 1 nucleotide of the D1 / D2 region.

```
                                10        20        30        40        50        60        70        80        90       100
                        ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.cerevisiae            TGGTACCTTCGGTGCCCGAGTTGTAATTTGGAGAGGGCAACTTTGGGGCCGTTCCTTGTCTATGTTCCTTGGAACAGGACGTCATAGAGGGTGAGAATCC
S.bayanus_CBS_380_[T]   .........T.....................................A....................................................
S.bayanus_CBS_395_[7]   .........T.....................................A....................................................

                               110       120       130       140       150       160       170       180       190       200
                        ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.cerevisiae            CGTGTGGCGAGGAGTGCGGTTCTTTGTAAAGTGCCTTCGAAGAGTCGAGTTGTTTGGGAATGCAGCTCTAAGTGGGTGGTAAATTCCATCTAAAGCTAAA
S.bayanus_CBS_380_[T]   .......................A............................................................................
S.bayanus_CBS_395_[7]   .......................A............................................................................

                               210       220       230       240       250       260       270       280       290       300
                        ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.cerevisiae            TATTGGCGAGAGACCGATAGCGAACAAGTACAGTGATGGAAAGATGAAAAGAACTTTGAAAAGAGAGTGAAAAAGTACGTGAAATTGTTGAAAGGGAAGG
S.bayanus_CBS_380_[T]   ....................................................................................................
S.bayanus_CBS_395_[7]   ....................................................................................................

                               310       320       330       340       350       360       370       380       390       400
                        ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.cerevisiae            GCATTTGATCAGACATGGTGTTTTGTGCCCTCTGCTCCTTGTGGGTAGGGGAATCTCGCATTTCACTGGGCCAGCATCAGTTTTGGTGGCAGGATAAATC
S.bayanus_CBS_380_[T]   .........................C....................G.............G.......................................
S.bayanus_CBS_395_[7]   .........................C....................G.............G.......................................

                               410       420       430       440       450       460       470       480       490       500
                        ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
S.cerevisiae            CATAGGAATGTAGCTTGCCTCGGTAAGTATTATAGCCTGTGGGAATACTGCCAGCTGGGACTGAGGACTGCGACGTAAGTCAAGGATGCTGGCATAATGG
S.bayanus_CBS_380_[T]   .G..........A.....T....G..............A..............A..............................................
S.bayanus_CBS_395_[7]   .G..........A.....T....G..............A.............................................................

                               510
                        ....|....|.
S.cerevisiae            TTATATGCCGC
S.bayanus_CBS_380_[T]   ...........
S.bayanus_CBS_395_[7]   ...........
```

Figure 5.30: ClustalW alignment of the D1 / D2 rDNA sequences of representative strains for each cluster shown in figure 5.29. *S. cerevisiae* represents the *S. cerevisiae*-like cluster. *S. bayanus* CBS 380 represents the *S. bayanus* "*bayanus*" subgroup and *S. bayanus* CBS 395 represents the *S. bayanus* "*uvarum*" subgroup. The single nucleotide difference between these two *S. bayanus* subgroups is highlighted at position 454. The *S. bayanus* CBS 380 and *S. cerevisiae* sequences were downloaded from the SGD website.

Species	CBS No.	NCYC No.	Local No.	Remarks
S. cerevisiae	CBS8803	-	-	Strain used for the *Saccharomyces* genome project (S288C)
S. bayanus T	CBS380	NCYC2578	-	Type strain
S. bayanus 1	CBS7001	NCYC2669	-	Former type strain of *S. abuliensis*
S. bayanus 7	CBS395	NCYC509	-	Former type strain of *S. uvarum*
S. paradoxus	CBS432	NCYC2600	B4239	Neotype strain
S. kudriavzevii	CBS8840	NCYC2889	-	Type strain
S. mikatae	CBS8839	NCYC2888	-	Type strain
S. cariocanus	CBS8841	NCYC2890	-	Type strain
S. pastorianus 1	CBS1538	NCYC392	-	Neotype strain
S. pastorianus 2	-	-	B4667	Formerly *S. bayanus var. uvarum*
S. pastorianus 3	-	-	B4417	Formerly *S. carlsbergensis*
S. pastorianus 4	CBS1513	NCYC396	B4243	Former type strain of *S. carlsbergensis*
S. pastorianus 5	CBS1260	NCYC400	B4242	-
S. pastorianus 6	CBS1503	NCYC2801	-	Former type strain of *S. monacensis*
S. pastorianus 7	-	NCYC453	-	Formerly *S. carlsbergensis*
S. pastorianus W-34/70	-	-	W-34/70	Weihenstephan strain

Table 5.10: *S. pastorianus* isolates included in this study and their origins and phylogenetic history.

The genes sequenced in the *S. pastorianus* and *S. bayanus* strains differed in their sequences. Table 5.11 shows the sequence identity (where 100% identical) or the sequence origin (where sequences are diverged) of each marker tested and some additional markers from the literature (*GDH1* and *MET2* [153]). The heterogeneity exhibited by the *S. pastorianus* strains suggests that species-specific PCR primers that would amplify all of them are implausible.

Species and strain	**CBS code**	**Marker**					
		D1 / D2	***FAL1***	***PRI1***	***MEX67***	***GDH1***	***MET2***
S. bayanus (T)	CBS 380	*Sb*	-	-	-	*Su**	*Su*
S. bayanus 1	CBS 7001	*Su*	*Su*	*Su*	*Su*	*Su*	*Su*
S. bayanus 7	CBS 395	*Su*	*Su*	*Su*	*Su*	*Su*	*Su*
S. pastorianus 1 (T)	CBS 1538	*Sb*	*Su*	*Su*	*Sp*	*Su**	*Sp*
S. pastorianus 2		*Su*	*Su*	*Su*	*Su*	-	-
S. pastorianus 3		*Sc*	*Sc*	*Sc*	*Sc*	-	-
S. pastorianus 4	CBS 1513	*Sb*	*Sp*[1]	*Sp*	*Sp*	*Su**	*Sp*
S. pastorianus 5		*Sc*	*Sp*[2]	*Sp*	-	-	-
S. pastorianus 6	CSB 1503	*Sb*	*Sp*[3]	*Sc**	*Sp*	*Su**	*Sp*
S. pastorianus 7		*Sc*	*Sp*[4]	*Sp*	-	-	-

Table 5.11: Sequence variation in *S. pastorianus* strains compared to *S. bayanus* strains.
(*Sb*) The sequence is identical to *S. bayanus* "bayanus" subgroup.
(*Su*) The sequence is identical to *S. bayanus* "uvarum" subgroup.
(*Su**) The sequence is not identical to *S. uvarum* but *S. uvarum* is the ancestor.
(*Sc*) The sequence is identical to *S. cerevisiae.*
(*Sc**) The sequence is not identical to *S. cerevisiae* but *S. cerevisiae* is the ancestor.
(*Sp*) The sequence is not like *S. uvarum* or *S. cerevisiae* but is identical to those designated Sp in other *S. pastorianus* strains (as the *S. bayanus* "bayanus" subgroup is not available for comparison of *MEX67*, it is possible that the *Sp* sequence is identical to the *S. bayanus* "bayanus" subgroup).
(*Sp*[1-4]) The sequence is not like *S. uvarum* or *S. cerevisiae* and is different in each of the other *S. pastorianus* strains.

5.2.4 Multiplexing

Saccharomyces species-specific primers were used in a single reaction mixture to test if they would still only amplify the correct DNA target without any false positive, false negative or primer dimerisation reactions. The results are shown in figures 5.31 and 5.32. When the target DNA is present, only a specific-sized amplicon is produced. When no target DNA is present, no reaction takes place despite the fact that 6 pairs of primers are present in the reagents. When the genomic DNA of two *Saccharomyces* species is present, two specific-sized amplicons are produced, although smaller amplicons are produced preferentially to larger ones. However, when the genomic DNA of one of each *Saccharomyces* species is present, only the smallest amplicons are produced.

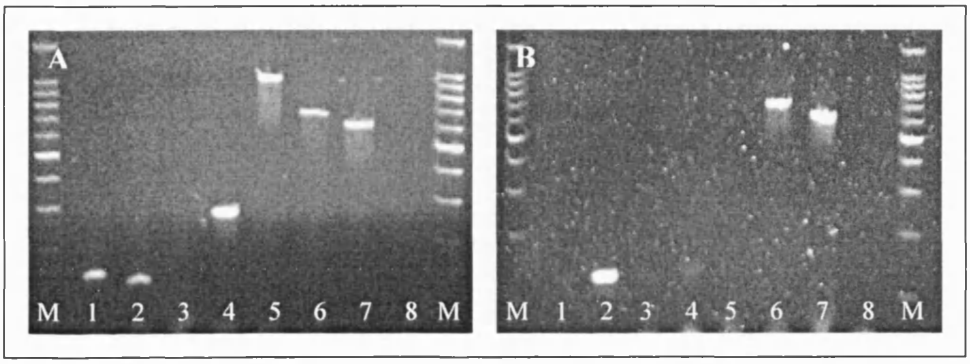

Figure 5.31: Multiplex PCR with (A) All *Saccharomyces* species-specific primers in a single reaction mixture tested against the genomic DNA of each *Saccharomyces* species (B) *S. bayanus*, *S. mikatae* and *S. kudriavzevii* species-specific primers in a single reaction mixture tested against the genomic DNA of each *Saccharomyces* species. DNA. Specific primer pairs used and amplicon size: *S. cerevisiae* F2 / R2 (150bp), *S. bayanus* F2 / R1 (128bp), *S. pastorianus* N/A, *S. paradoxus* F1 / R2 (265bp), *S. cariocanus* F1 / R1 (1016bp), *S. mikatae* F1 / R1 (707bp) and *S. kudriavzevii* F2 / R1 (660bp). (1) *S. cerevisiae* 1, (2) *S. bayanus* 2, (3) *S. pastorianus* 1, (4) *S. paradoxus* 1, (5) *S. cariocanus* 1, (6) *S. mikatae* 1, (7) *S. kudriavzevii* 1, (8) No DNA

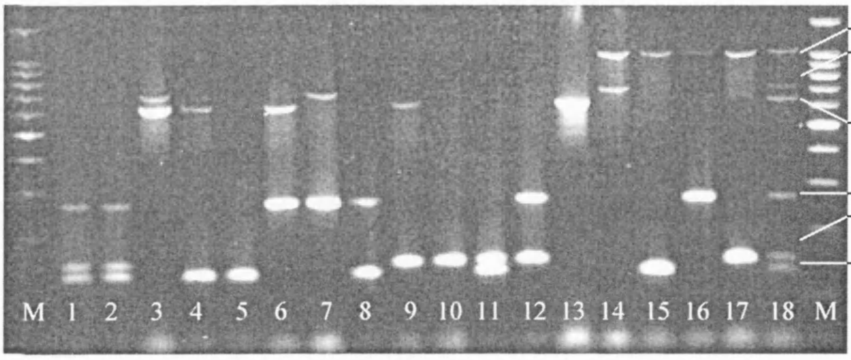

Figure 5.32: Multiplex PCR with all *Saccharomyces* species-specific primers in a single reaction mixture, tested against different combinations of *Saccharomyces* species genomic DNA.

Specific primers used and amplicon size: *S. cerevisiae* F2 / R2 (150bp), *S. bayanus* F2 / R1 (128bp), *S. paradoxus* F1 / R2 (265bp), *S. cariocanus* F1 / R1 (1016bp), *S. mikatae* F1 / R1 (707bp) and *S. kudriavzevii* F2 / R1 (660bp).

Genomic DNA in each lane: (1) All species, (2) All species, (3) *S. mikatae* 1 / *S. kudriavzevii* 1, (4) *S. bayanus* 2 / *S. kudriavzevii* 1, (5) *S. bayanus* 2 / *S. mikatae* 1, (6) *S. paradoxus* 1 / *S. kudriavzevii* 1, (7) *S. paradoxus* 1 / *S. mikatae* 1, (8) *S. paradoxus* 1 / *S. bayanus* 2, (9) *S. cerevisiae* 1 / *S. kudriavzevii* 1, (10) *S. cerevisiae* 1 / *S. mikatae* 1, (11) *S. cerevisiae* 1 / *S. bayanus* 2, (12) *S. cerevisiae* 1 / *S. paradoxus* 1, (13) *S. cariocanus* 1 / *S. kudriavzevii* 1, (14) *S. cariocanus* 1 / *S. mikatae* 1, (15) *S. cariocanus* 1 / *S. bayanus* 2, (16) *S. cariocanus* 1 / *S. paradoxus* 1, (17) *S. cariocanus* 1 / *S. cerevisiae* 1, (18) Species ladder (Negative control result not shown, no amplification).

5.3 Discussion

5.3.1 D1/D2 sequences

The D1 / D2 sequences of the *Saccharomyces* species are highly related (figure 5.1). The maximum number of differences occurs between *S. cerevisiae* and *S. mikatae*, 14 SNPs occur over 500bp of sequence. The minimum number of differences occurs between *S. paradoxus* and *S. cariocanus* with only 1 SNP between the two species over the whole region. The sequence of the *S. pastorianus* type strain (CBS 392) is identical to the *S. bayanus* type strain (CBS 380). *S. pastorianus* strains are allopolyploid and so distinguishing them from the parental *S. cerevisiae* and *S. bayanus* species is difficult. Discounting these strains, *S. paradoxus* and *S. cariocanus* are the most closely related siblings represented in this project and as such, should be the most difficult to distinguish between with oligonucleotide primers.

5.3.2 Specific primer design

Designing species-specific primers for the individual *Saccharomyces* species demonstrated how much more simple and straightforward the technique can be when applied to a well-defined and extensively studied clade. The amount of sequence data available from comprehensive genome surveys allowed a single gene sequence to be tailored for suitability to each species. For instance, the *S. kudriavzevii* specific reverse primer was targeted to a locus where the other species had a 12-nucleotide (4 amino acid) insertion in the sequence (figure 5.13). The *S. bayanus* specific forward primer was targeted to a locus in which the other species had various numbers of codon insertions (figure 5.25). The unsuccessful *S. pastorianus* forward primers were designed to target what was thought to be a species-specific codon insertion.

The sequence data was not required for all species of the genus but it is essential to have it for the sibling of the target species. *S. cerevisiae* primers were designed successfully without *S. kudriavzevii* sequence data. This shows that providing the closest sibling sequence is available, species-specific primers are still feasible. The species from which sequence data is missing must be practically tested to prove that the primer target locus is sufficiently different to prevent primer annealing. Using a single gene per species also increases the likelihood that a multiplex approach will work for identifying species of the genus.

5.3.3 Specific primer testing

Primer testing provided similar results to the other genera examined. Only one species required additional optimisation. *S. paradoxus* primers amplified *S. cariocanus* isolates under the generic reaction conditions and were only specific at an annealing temperature of 65°C. *S. bayanus* primers were not shown to be specific as they all amplified *S. pastorianus* strain 2 but no other strains or species.

Some primer pairs annealed to unexpected regions. *S. paradoxus* strains produced an oversized weak product with *S. cerevisiae* F1 / R1 primers. S *kudriavzevii* primer pair F1 / R2 produced a smaller than expected product with all species, particularly strong in *S. cerevisiae*. *S. bayanus* F1 / R1 produced a massively oversized weak product with *S. cerevisiae* isolates. *S. bayanus* F1 / R2 produced an oversized weak product with *S. paradoxus* strains. Fungal WU BLAST version 2.0 against the *S. paradoxus* genome on the *Saccharomyces* Genome Database website [154] did not clarify where the primers could be binding to produce these unexpected amplicons. NCBI BLASTn [126] against the Fungal DNA database and *Saccharomyces cerevisiae* complete genome had similar results.

5.3.4 *S. cariocanus / S. paradoxus*

With only one SNP in the whole D1 / D2 region, *S. cariocanus* and *S. paradoxus* are closely related which would explain the mating experiment results reported in Sniegowski *et al* (2002) [158]. The confusion regarding the identity of the *S. cariocanus* isolates was highlighted whilst testing the *S. paradoxus* specific primers (figure 5.20). D1 / D2 sequence analysis and *S. cariocanus* specific primers resolved this confusion and a higher annealing temperature was required for the *S. paradoxus* specific primers to be successful.

It has recently been claimed that *S. cariocanus* and *S. paradoxus* are the same species (Ed Louis, personal communication) and that *S. cariocanus* strains are American variants of *S. paradoxus*. *S. cariocanus* strains 1 and 2 originate from South America while strains 3 – 11 originate from the USA.

This confusion between species identity highlights the danger of relying too heavily on D1 / D2 sequence data for species identification. Some species have been shown to have unusually polymorphic D1 / D2 regions [159]. Sequence data provides a useful guideline for determining species boundaries, however the only definitive tool is using mating reactions and the biological species concept.

5.3.5 *S. pastorianus*

The *S. pastorianus* strains tested with the potentially *S. pastorianus* specific primers gave different results. Strains 4 and 5 produced the desired amplicon. Strains 1 – 3 did not. The *S. pastorianus* strains tested with the specific primers of other species gave a variety of results. The most common result was no amplification, but they were amplified with a few primer pairs (results are summarised in table 5.12). This showed that at these primer loci, the gene sequences differ among the *S. pastorianus* strains.

As the D1 / D2 sequences of *S. pastorianus* strains were identical to either *S. bayanus* CBS 380, *S. bayanus* CBS 395 or *S. cerevisiae* it is clear that this region cannot be used for the identification of *S. pastorianus* species. As the results with different primer pairs show that the *S. pastorianus*, *S. cerevisiae* and *S. bayanus* strains behave differently, this is not a simple problem of incorrect identification as was the case with *S. cariocanus* and *S. paradoxus*.

Species and strain	**Primer pairs and gene target**		
	All *S. bayanus* primers (*DBP6*)	*S. cariocanus* F1 / R2 (*VAS1*)	*S. cariocanus* F3 / R3 (*VAS1*)
S. bayanus 1 – 6	✓	×	✓
S. cerevisiae 1 – 8	×	×	✓
S. pastorianus 1	×	×	×
S. pastorianus 2	✓	×	✓
S. pastorianus 3	×	×	✓
S. pastorianus 4	×	✓	×
S. pastorianus 5	×	×	×

Table 5.12: A comparison of the PCR results using different primer pairs on *S. bayanus*, *S. cerevisiae* and *S. pastorianus* strains. (✓) Amplification. (×) No amplification. Refer to table 5.9 for strain histories.

The *FAL1*, *MEX67* and *PRI1* gene sequences were experimentally obtained from as many *S. pastorianus* strains as possible and then aligned using ClustalW. Some of the markers in some of the strains were identical to either *S. bayanus* 7 (*S. bayanus* "*uvarum*" subgroup) or *S. cerevisiae*. Some had the sequence from *S. bayanus* 7 or *S. cerevisiae* as a recent ancestor. Some sequences were suitably diverged from *S. bayanus* 7 and/or *S. cerevisiae* to be considered "*S. pastorianus*" sequences (table 5.11). Single gene phylogenetic trees have been produced for each of the 3 marker genes (figures 5.33 to 5.35). They illustrate the similarities in some of the sequences.

Caseregola *et al* (2001) [152] investigated the beer yeast genome to establish what species or strain was the second *S. pastorianus* parent along with *S. cerevisiae*. They concluded that *S. bayanus* strains of the "*bayanus*" subgroup (represented by CBS 380, CBS 1501 and CBS 425) are good candidates to have been involved in the formation of modern lager hybrids.

The *S. bayanus* type strain (CBS 380) seems to have undergone the exchange of genetic material with *S. bayanus* CBS 395 of the *S. bayanus* "*uvarum*" subgroup which is why it shares the *MET2* and an ancestral *GDH1* sequence with this strain whilst having a "*bayanus*" D1 / D2 sequence.

All of the markers tested in *S. pastorianus* strain 2 are identical to *S. bayanus* subgroup "*uvarum*". This strain was amplified with all of the *S. bayanus* specific primer pairs. It is plausible that this strain should be reclassified as *S. bayanus*, although this would not be clear without karyotype analysis and the investigation of more genetic markers. If it was found to be an *S. bayanus* strain, then the *S. bayanus* specific primers would indeed be species-specific as strain 2 was the only other *Saccharomyces* that they amplified.

All of the markers tested in *S. pastorianus* strain 3 are identical to *S. cerevisiae*. Similarly to above, further investigation would be required to establish if it is in fact *S. pastorianus* or *S. cerevisiae*.

The remaining *S. pastorianus* strains seem to have different origins. The gene sequences examined do not have a clear lineage to link them all back to a single rare mating between an *S. bayanus* and *S. cerevisiae* strain. For example, the *PRI1* gene has "*uvarum*" origin in *S. pastorianus* 1, while *S. pastorianus* 4, 5, 6 and 7 share an identical sequence that does not obviously originate from either "*uvarum*" or *S. cerevisiae*. The *FAL1* gene has "*uvarum*" origin in *S. pastorianus* 1, while 4, 5, 6 and 7 all have unique sequences. On this evidence, it would seem that the *S. pastorianus* type strain (strain 1, CBS 1538) has an *S. bayanus* subgroup "*uvarum*" strain as a parent. However, the *MEX67* sequence is a unique *S. pastorianus* sequence, the *GDH1* sequence is a divergent form of the "*uvarum*" sequence and the *MET2* gene is identical to "*bayanus*" sequence. This result highlights the necessity of testing multiple markers as a small number can give false impressions.

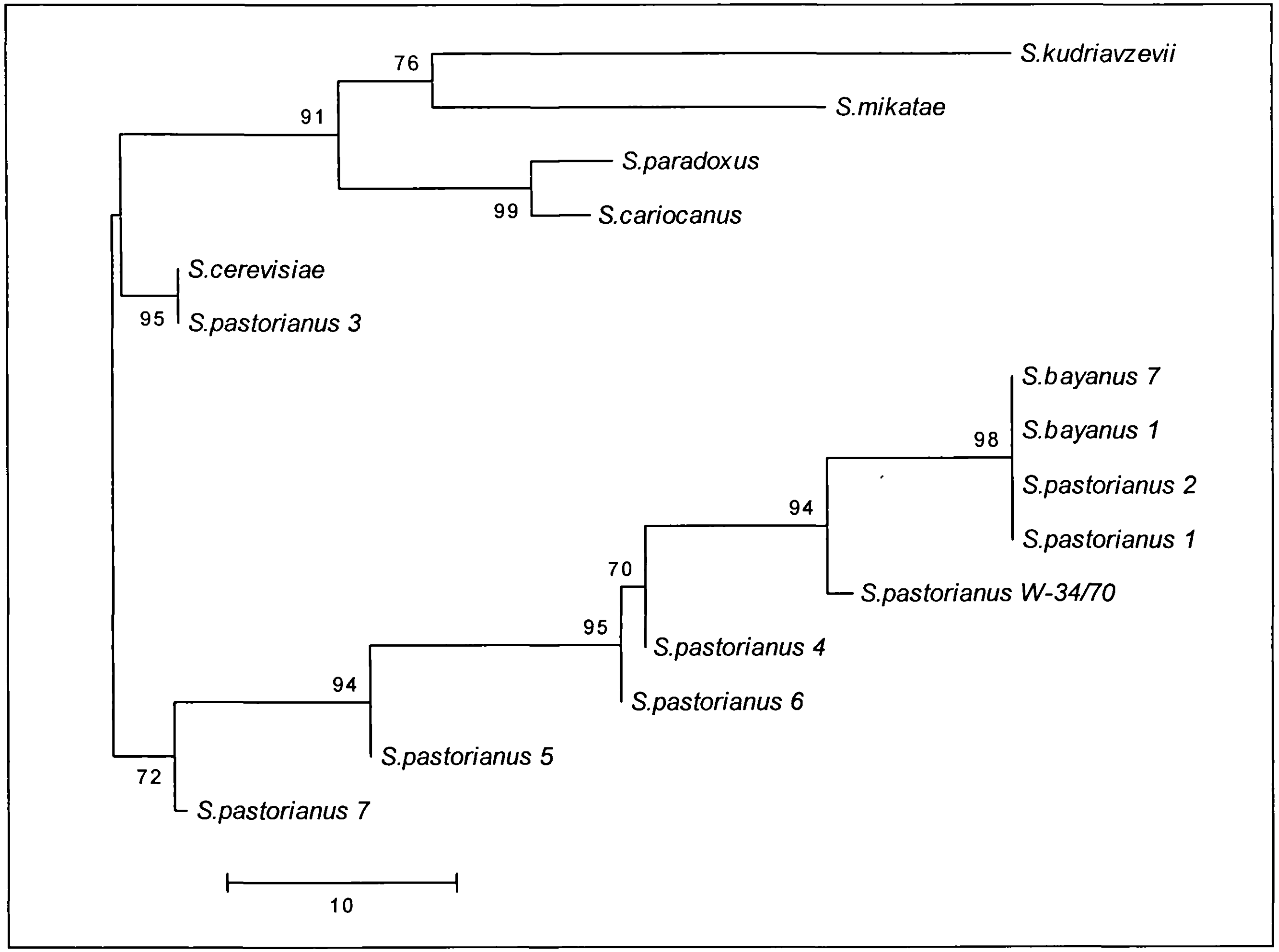

Figure 5.33: Maximum Parsimony phylogenetic tree of the *Saccharomyces* type species (barring *S. bayanus*), *S. bayanus* strains and *S. pastorianus* strains *FAL1* gene. Number of bootstrap replicates = 100.

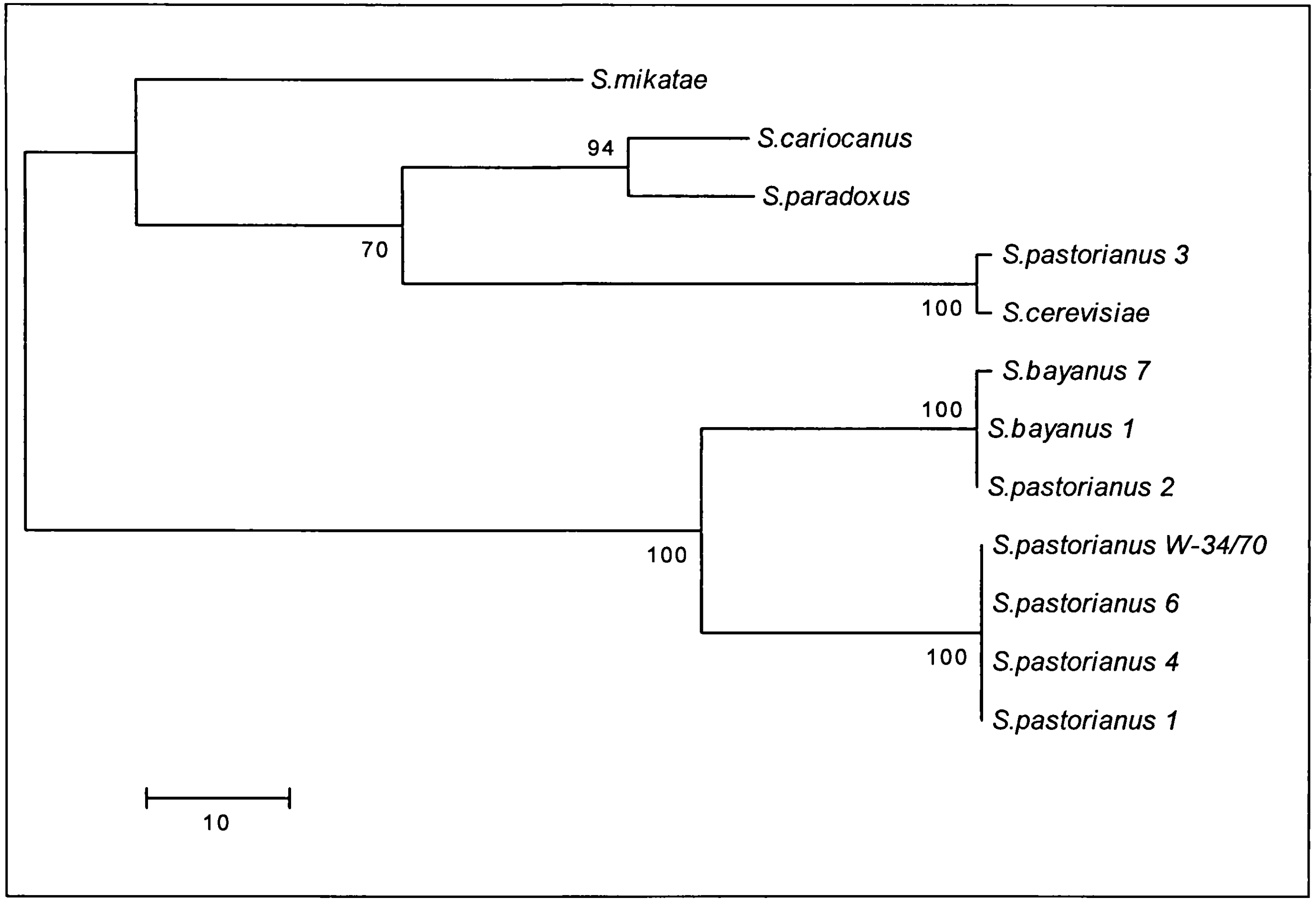

Figure 5.34: Maximum Parsimony phylogenetic tree of the *Saccharomyces* type species (barring *S. bayanus*), *S. bayanus* strains and *S. pastorianus* strains *MEX67* gene. Number of bootstrap replicates = 100.

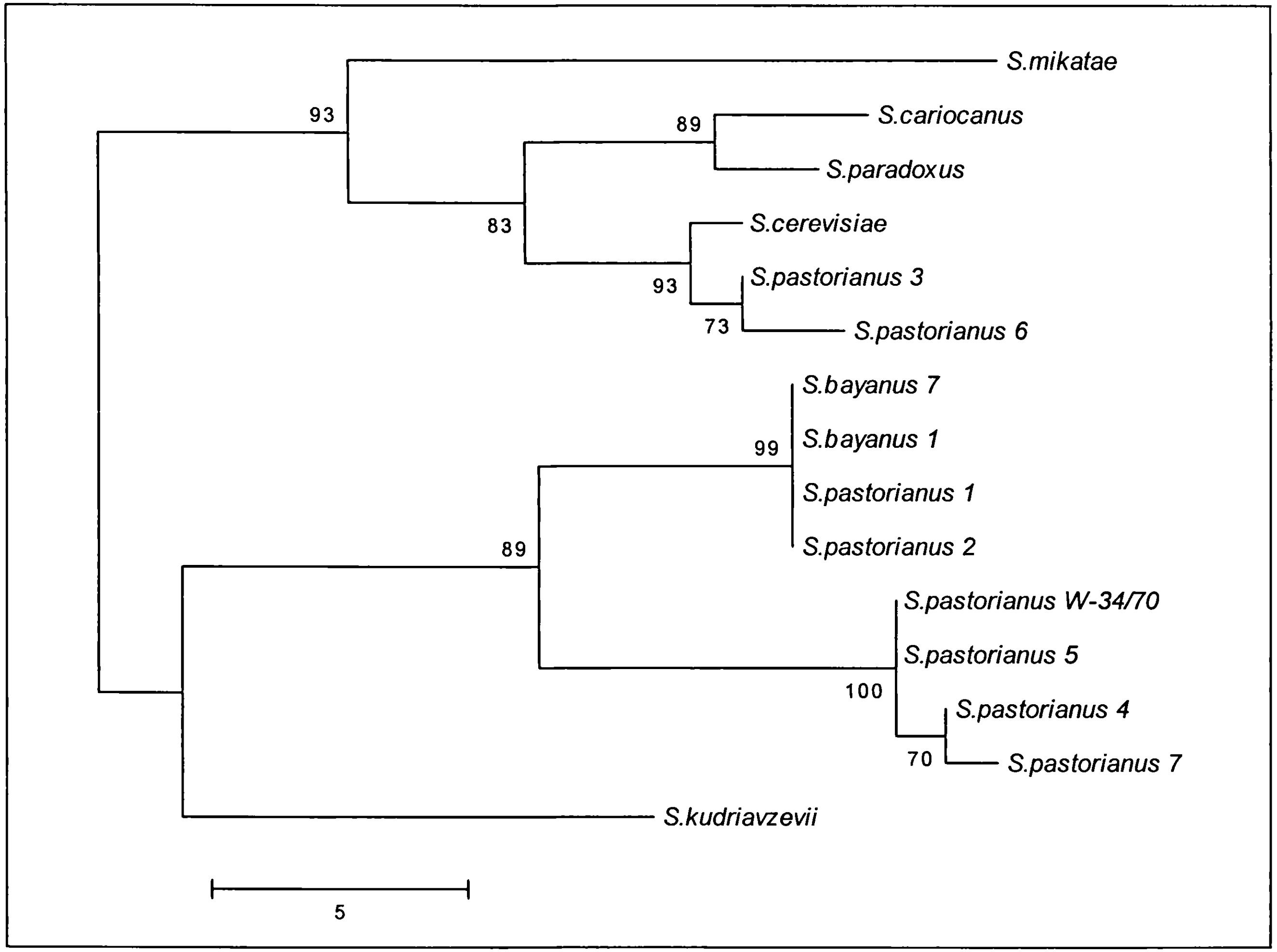

Figure 5.35: Maximum Parsimony phylogenetic tree of the *Saccharomyces* type species (barring *S. bayanus*), *S. bayanus* strains and *S. pastorianus* strains *PRI1* gene. Number of bootstrap replicates = 100.

The lack of a *FAL1*, *MEX67* and *PRI1* sequence from a *S. bayanus* strain of the "*bayanus*" subgroup leaves it unclear if non-*cerevisiae* and -*uvarum* sequences of these genes are truly unique *S. pastorianus* sequences or could originate from this possible ancestor. The basic topology of the *MEX67* and *PRI1* trees (figures 5.34 and 5.35) are similar to the D1 / D2 tree topology (figure 5.29) which does include *S. bayanus* CBS 380.

S. pastorianus is reported to be an allopolyploid species, containing parts of two divergent genomes. Casaregola *et al* [152] found alleles of the same gene that differed in parent species in *S. pastorianus* strains 4 (CBS 1513, formerly *S. carlsbergensis*) and 6 (CBS 1503, formerly *S. monacensis*). The genes were *HIS4* and *URA3*. When the *FAL1*, *MEX67* and *PRI1* sequences were amplified from these strains, there was no evidence of multiple alleles. Each PCR reaction produced a single amplicon and sequencing results were clear and without double peaks indicating the presence of only one sequences. The degenerate primers that amplified the genes from *S. pastorianus* 1 – 7 were designed based on an alignment of *S. pastorianus* W-34/70 (which seems to be non-"*uvarum*" like), *S. bayanus* 7 ("*uvarum*" subgroup) and *S. cerevisiae*. The fact that they amplified unique, *S. cerevisiae*-like and "*uvarum*"-like sequences suggests no inherent bias towards one type of allele. In the study by Nguyen *et al* [150] investigating the *GDH1* and *MET2* sequences in strains 1, 4 and 6, no multiple alleles were found.

It would appear that prior to the mating event that resulted in hybridisation with *S. cerevisiae*, the *S. bayanus* parent strains had a varying number of genes from the "*uvarum*" or "*bayanus*" subgroups depending on individual incidents of genetic exchange with each other. These heterogeneous genomes then contributed to hybrid *S. pastorianus* strains. In a brewing or wine making environment, multiple species would be concentrated in a nutrient rich environment that would potentially facilitate genetic exchange between closely related organisms.

On this evidence, species-specific primers would be impossible to design for *S. pastorianus*. If each *S. pastorianus* strain has originated from a different hybridisation event involving *S. cerevisiae* and a heterogeneous group of *S. bayanus* strains, they will not form a discrete homogeneous species that could be classified under the umbrella of a single taxon. This also has implications upon the specificity of *S. bayanus* and *S. cerevisiae* primers. Although these primers have been shown to be able to amplify all target species, in a brewing environment where all three species are

The lack of a *FAL1*, *MEX67* and *PRI1* sequence from a *S. bayanus* strain of the "*bayanus*" subgroup leaves it unclear if non-*cerevisiae* and -*uvarum* sequences of these genes are truly unique *S. pastorianus* sequences or could originate from this possible ancestor. The basic topology of the *MEX67* and *PRI1* trees (figures 5.34 and 5.35) are similar to the D1 / D2 tree topology (figure 5.29) which does include *S. bayanus* CBS 380.

S. pastorianus is reported to be an allopolyploid species, containing parts of two divergent genomes. Casaregola *et al* [152] found alleles of the same gene that differed in parent species in *S. pastorianus* strains 4 (CBS 1513, formerly *S. carlsbergensis*) and 6 (CBS 1503, formerly *S. monacensis*). The genes were *HIS4* and *URA3*. When the *FAL1*, *MEX67* and *PRI1* sequences were amplified from these strains, there was no evidence of multiple alleles. Each PCR reaction produced a single amplicon and sequencing results were clear and without double peaks indicating the presence of only one sequences. The degenerate primers that amplified the genes from *S. pastorianus* 1 – 7 were designed based on an alignment of *S. pastorianus* W-34/70 (which seems to be non-"*uvarum*" like), *S. bayanus* 7 ("*uvarum*" subgroup) and *S. cerevisiae*. The fact that they amplified unique, *S. cerevisiae*-like and "*uvarum*"-like sequences suggests no inherent bias towards one type of allele. In the study by Nguyen *et al* [150] investigating the *GDH1* and *MET2* sequences in strains 1, 4 and 6, no multiple alleles were found.

It would appear that prior to the mating event that resulted in hybridicity with *S. cerevisiae*, the *S. bayanus* parent strains had a varying number of genes from the "*uvarum*" or "*bayanus*" subgroups depending on individual incidents of genetic exchange with each other. These heterogeneous genomes then contributed to hybrid *S. pastorianus* strains. In a brewing or wine making environment, multiple species would be concentrated in a nutrient rich environment that would potentially facilitate genetic exchange between closely related organisms.

On this evidence, species-specific primers would be impossible to design for *S. pastorianus*. If each *S. pastorianus* strain has originated from a different hybridisation event involving *S. cerevisiae* and a heterogeneous group of *S. bayanus* strains, they will not form a discrete homogeneous species that could be classified under the umbrella of a single taxon. This also has implications upon the specificity of *S. bayanus* and *S. cerevisiae* primers. Although these primers have been shown to be able to amplify all target species, in a brewing environment where all three species are mixed, there is a likelihood of *S. pastorianus* strains occurring with a *bayanus*- or *uvarum*-like *DBP6* or a *cerevisiae*-like *MEX67* gene that would be amplified. In terms of food and drink spoilage, *Saccharomyces* are the main contaminants in plants producing both beer and fruit juices. The implications on attempts to identify *S. cerevisiae* and *S. bayanus* are that if an amplicon is produced, the causative species could be the target or *S. pastorianus*. If an amplicon is not produced, *S. pastorianus* may still be present.

5.3.6 Multiplexing

Multiplex experiments that have all of the species-specific primer pairs included in a single reaction mixture were successful (figures 5.31 and 5.32). When the target species genomic DNA was not present, the species-specific product was not detected on the agarose gel. When the target species genomic DNA was present, a species-specific amplicon was produced. This was also successful when two different target genomes were present and when no target genome was present. However, when the genomic DNA for all six-target species was present (*S. pastorianus* was not included), only the smallest three amplicons were produced. This is because a higher frequency of complete polymerisation reactions can take place with smaller amplicons producing a high ratio of smaller amplicons to large amplicons in the resulting reaction mixture. Whilst the amount of reagents could be increased to ensure that all six specific amplicons would be produced, this would alter the dynamics of the reaction and so primer specificity could not be guaranteed. The likelihood of more than one species being involved in a spoilage event is low enough that this is not a problem.

5.4 Conclusion

Species-specific PCR primers have been successfully designed for *S. cariocanus*, *S. kudriavzevii*, *S. mikatae* and *S. paradoxus*. *S. cariocanus* and *S. paradoxus* have since been merged into a single taxa but the *S. paradoxus* primers are specific for all variants at a lower annealing temperature. *S. bayanus* and *S. cerevisiae* primers exclude all other species barring *S. pastorianus*. The extent of amplification of *S. pastorianus* strains with these primers cannot be determined while the extent of the hybrid nature of this taxon remains unknown. *S. pastorianus* specific primers cannot be successfully designed for the same reason.

Using a single target gene per species method, reactions can be multiplexed thus reducing the number of individual tests required. This improved the speed and efficiency of a rapid, high throughput species identification method.

Chapter 6. Concluding discussion

6.1 Primer characteristics

The primer target loci of the *Zygosaccharomyces* primers have been examined for differences in mismatch numbers, melting temperatures and 3' mismatches. Each primer had 6 target loci, one in each of the *Zygosaccharomyces* species. There were 74 primer pairs tested, a total of 444. When each primer pair was initially screened on pairs of species, the primers that amplified both or neither of the pairs were scored and included in these results. The primers that only amplified one member of a pair of strains have not been included, as the binding was judged to be weak and variable. This left 292 data points to examine.

6.1.1 Total number of mismatches

When designing primers that are tailored to be species-specific, maximising the number of mismatches in the non-target genome is the main priority. If the non-target species can be eliminated, sub-optimal reaction conditions can be used to ensure that even strains of heterogeneous taxa can be identified with a specific primer pair. As the number of mismatches increased in the *Zygosaccharomyces* primer targets, the proportion of non-specific results decreased (figure 6.1).

6.1.2 3' mismatches

It has been reported that the 3' base of a primer is the most important in primer binding and successful polymerisation initiation [160, 161, 162]. Due to this, one of the primer design parameters included incorporating a 3' mismatch into as many primers as possible. The results for the specificity of primer pairs with two, one or no 3' mismatches were analysed. When both primers in a pair had 3' mismatches, the chances of a negative reaction were not enhanced (table 6.1). A chi-square test (χ^2) comparing the proportions gives a P-value > 0.9 (χ^2 = 0.154 with 2 degrees of freedom). This shows that there is insufficient evidence to conclude that the proportion of positive and negative results differs when two, one or no 3' mismatches are present.

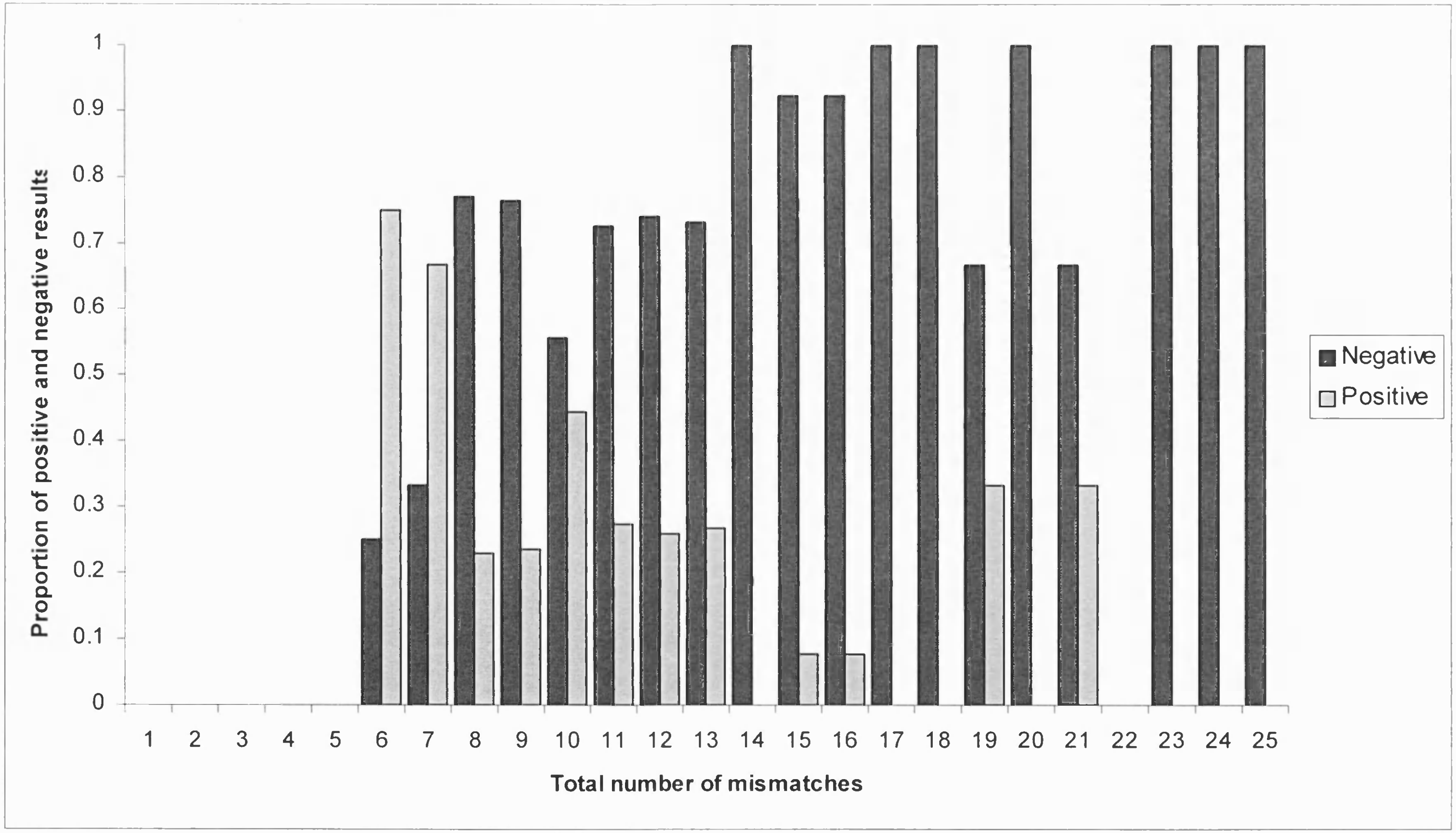

Figure 6.1: The total number of mismatches in the target loci of a primer pair compared to the proportion of positive and negative results with non-target species. Total number of primer results examined = 292. The maximum number of mismatches was 25 and the minimum number was 6. As mismatch number increases, the proportion of positive results decreases.

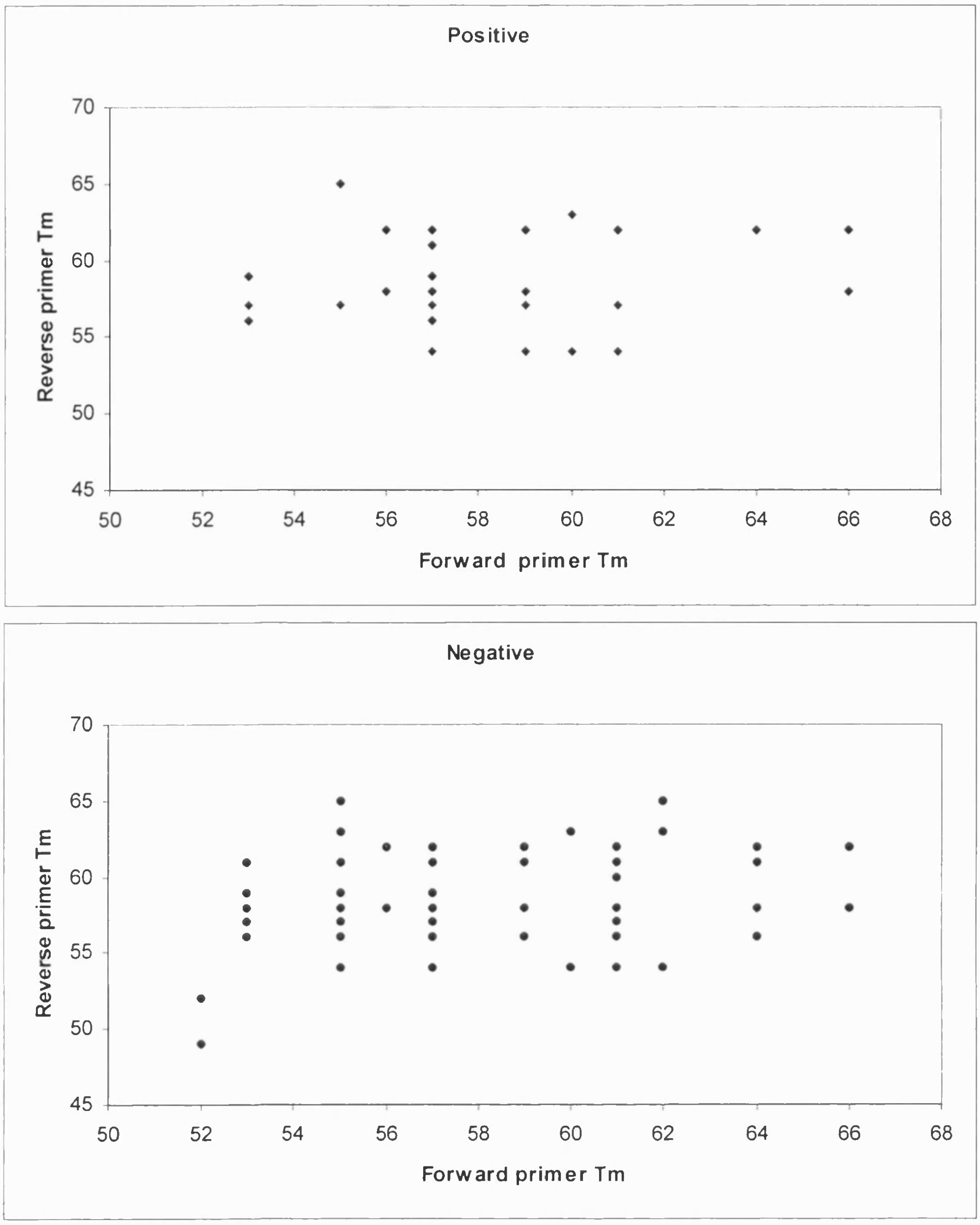

Figure 6.2: The melting temperature (°C) of the forward and reverse *Zygosaccharomyces* primers separated for positive and negative reactions. There is no trend in the plots to suggest that melting temperature had an affect on the specificity of the primers tested at 55°C.

Number of 3' mismatches	Positive results	% of total positive	Negative results	% of total negative	Total
2	10	16.9	35	19.1	45
1	28	47.5	84	45.9	112
0	21	35.6	64	35	85

Table 6.1: The number of positive and negative results with non-target species when the 3' position is mismatched in both primers of a pair, one primer and no primers.

6.1.3 Primer melting temperatures

The melting temperatures of primers are crucial in primer binding. However, as the primers were designed to amplify the true targets under sub-optimal conditions, the melting temperature did not have an affect on the outcome of specific or non-specific reactions (figure 6.2). However, the melting temperature should still be designed to be as high as possible in the eventuality that primers require temperature optimisation to improve their specificity. When a primer can bind to the true target at higher temperatures, there is more scope for the elimination of non-specific annealing.

Thermodynamic profiles for each possible individual mismatch within primers have been established [163, 164, 165, 166]. Although basic inferences could be made, the thermodynamic stability of each mismatch is context dependent i.e. it depends on the Watson-Crick closing pairs on each side of a polymorphism. Due to this, the potential effect of multiple mismatches on the thermodynamic character of a primer is difficult to determine. As 3' mismatches are deemed the most crucial, the effect upon melting temperature of the most 3' mismatch in the *Z. bailii* specific primer pair F1 / R2 was calculated [167] in 4 different yeast strains (*Z. bailii* strains 1 and 2, *Z. rouxii*, *Z. lentus* and *Z. bisporus*) that gave different results (table 6.2). The percentage of primer bound at different annealing temperatures illustrate that when multiple mismatches are involved, the 3' most polymorphism does not have the determining effect upon primer annealing and amplification (figures 6.3 and 6.4). The *Z. bailii* F2 / R1 primer combination did not amplify *Z. bisporus* at 55°C or higher, but the percentage bound at all temperatures is higher than that of *Z. bailii* 2, which was faithfully amplified up to 69.9°C. Taking the reverse primer into account does not resolve the matter as the *Z. bisporus* 3' mismatch is equal to *Z. rouxii* and highly similar to *Z. bailii* 2 and *Z. lentus* in the effect on primer binding.

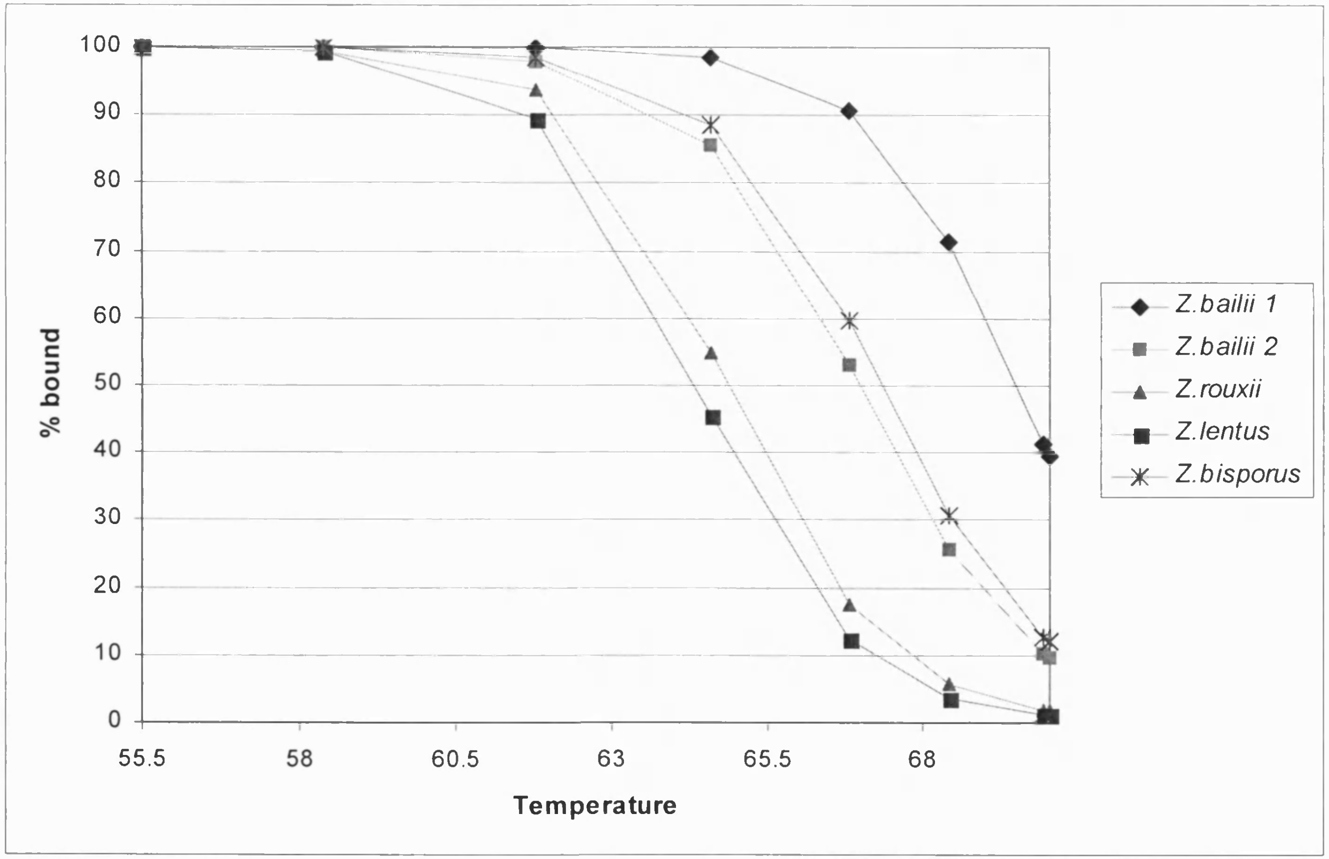

Figure 6.3: The percentage of *Z. bailii* F1 primer bound to different targets at the experimental annealing temperatures displayed in figure 4.19. *Z. bailii* strain 1 is amplified at all temperatures, *Z. bailii* strain 2 amplification is lost at 70°C, *Z. rouxii* is only amplified at 55.5°C, *Z. lentus* amplification persists until 69.9°C and *Z. bisporus* is not amplified at all (result not shown). The primer target loci are illustrated in figure 4.32. The most 3' mismatches are as follows (numbers derived from counting 3' position as 1 and increasing towards the 5'): *Z. bailii* 1 – No mismatch, *Z. bailii* 2 – Position 10 C – A, *Z. rouxii* – Position 7 G – A, *Z. lentus* – Position 7 G – G and *Z. bisporus* – Position 4 T – C

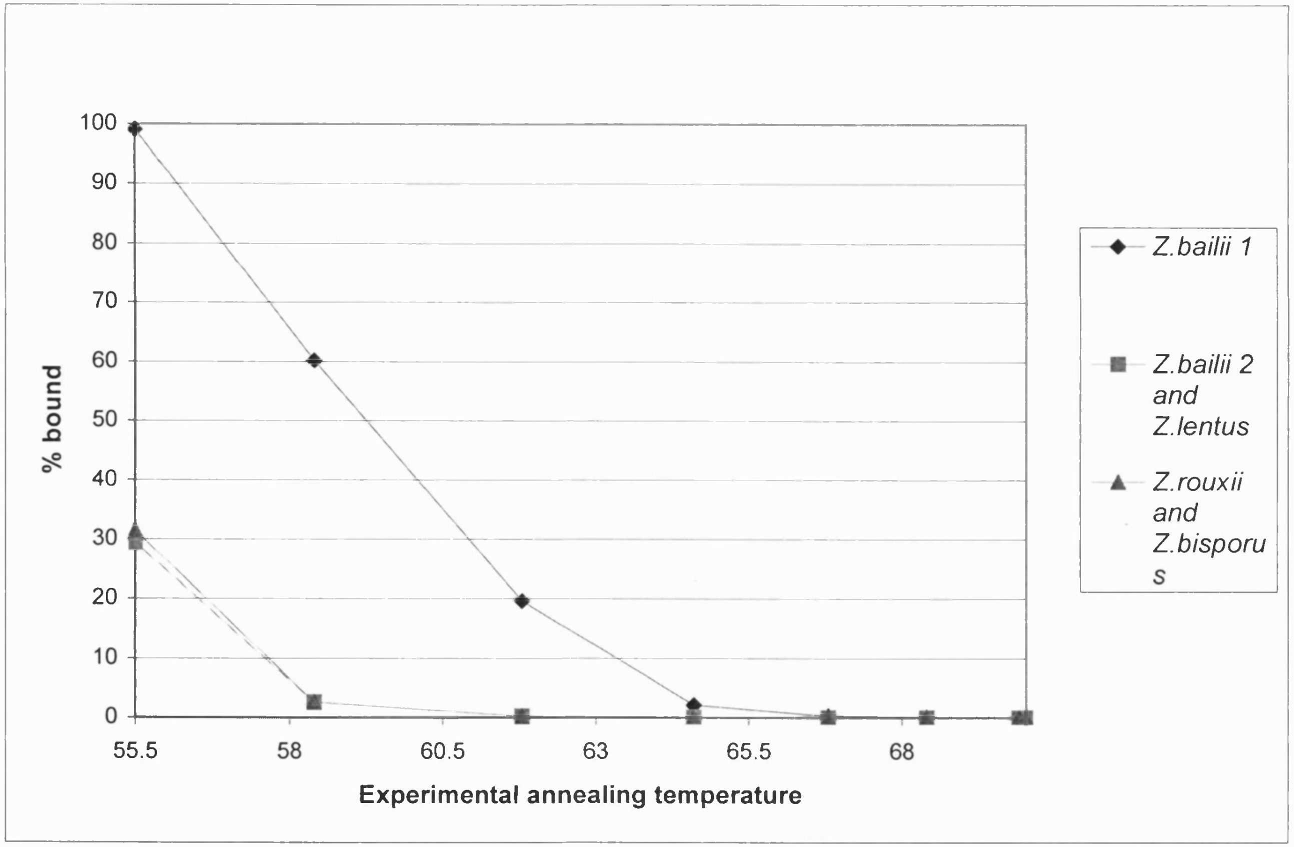

Figure 6.4: The percentage of *Z. bailii* R2 primer bound to different targets at the experimental annealing temperatures displayed in figure 4.19. *Z. bailii* strain 1 is amplified at all temperatures, *Z. bailii* strain 2 amplification is lost at 70°C, *Z. rouxii* is only amplified at 55.5°C, *Z. lentus* amplification persists until 69.9°C and *Z. bisporus* is not amplified at all (result not shown). The primer target loci are illustrated in figure 4.32. The most 3' mismatches are as follows (numbers derived from counting 3' position as 1 and increasing towards the 5'): *Z. bailii* 1 – No mismatch, *Z. bailii* 2 – Position 2 C – A, *Z. rouxii* – Position 2 C – T, *Z. lentus* – Position 2 C – A and *Z. bisporus* – Position 2 C – T

Species / strains	Reaction result	Total number of mismatches in *Z. bailii* F1 / R2 primer target loci
Z. bailii 1	Amplified up to 70°C	0
Z. bailii 2	Amplified up to 69.9°C	4
Z. bisporus	Not amplified	10
Z. lentus	Amplified up to 68.4°C	6
Z. rouxii	Amplified up to 55.5°C	10

Table 6.2: The PCR results using the *Z. bailii* F1 / R2 primer pair and the total number of mismatches in the specific and non-specific target loci.

These results show that when reactions conditions are sub-optimal, the most important factor in species-specific primer design is the total number of mismatches with the non-target species.

6.2 Genus specific primers

As well as species-specific primers, it should be possible to design primers that are able to amplify all members of a genus. Using the *Zygosaccharomyces HIS3* gene as an example, figure 6.5 shows that there are at least 2 regions where primers could be designed that have more mismatches in the outgroup species, *L. cidri*, than in the target species.

6.3 Conclusion

The rapid identification of yeasts using a species-specific PCR primer method has been illustrated. When designing primers, it is essential to maximise the number of mismatches with non-target species within the sequence. Mismatches with the closest sibling species are of paramount importance. When multiple gene sequences are available for species of a genus, the gene sequence used can be tailored towards a particular species depending on the amount of polymorphism. As more complete genomes are sequenced, species-specific genes may be discovered and exploited.

The genomic DNA sequence used as a primer target is dependent on the degree of similarity between the target species. Housekeeping gene sequences are suitable for closely related species such as *Zygosaccharomyces spp.* and *Saccharomyces spp*. For more diverse genera like *Brettanomyces / Dekkera*, rDNA is a suitable target for specific primers.

This method is only possible on well-examined genera that have had their species relationships completely elucidated and have been shown to exist as well-defined, monophyletic taxa. The D1 / D2 sequence data should not be relied upon too heavily when determining species relationships. However, all populations are variable and the sub-optimal nature of the reaction conditions allows the identification of all members of heterogeneous taxa.

Hybrid strains are difficult to identify using this technique. Conclusions cannot be drawn until the exact nature of any hybrids have been determined, as well as their relationships with parental species and siblings.

New species do not compromise the method although all primers of the target genus would have to be re-examined to ensure they did not amplify the new species. Primers to identify a new species would then need to be tested against all other members of the genus. Altering the taxonomy of existing species by fusing them into a single taxon is not a problem. The specific primers of sibling species have already been shown to be negative. A new primer pair would need to be chosen from existing pairs or new primers would need to be designed.

Using rapid DNA extraction, yeast species can be identified from a colony within 3 hours. There is potential to further increase the rapidity of the technique using a gel free system and real-time PCR. Multi-well microtitre plates can be used instead of individual tubes to increase throughput and allow potential for automation.

New sensitive techniques for the rapid identification of yeast and fungal species from a variety of substrates are now being investigated [92, 93, 94, 168]. Species can be identified and quantified in a single method or combination of methods. However, they have not been shown to be generic methods with potential for use in a range of taxa. The aim of this project was to produce a set of species-specific PCR primers that would rapidly and reliably identify any target yeast using a simple and low-cost method. This method has been successfully established and applied to the identification of food spoilage yeasts.

```
                             10        20        30        40        50        60        70        80        90       100
                     ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii HIS3        GCGAAGCACTCGGGTTGGTCTTTAATAGTCGAATGCATCGGCGATTTGCATATTGATGACCACCACACCACTGAGGATTGTGGCATTGCTCTCGGTGAAG
Z.bisporus HIS3      .....A.....C..G........G..C.....G.....T...........C..C..C..............A.....C.....T.....G..A.......
Z.lentus HIS3        ...........C........A..G..C..G..G..T...............C........T..............A..C.....T.....C..T.......
Z.kombuchaensis HIS3 ...........C........A..G..C.....G..T..T..T...........C.....T........T.....A..C.....T.....C..T.......
Z.rouxii HIS3        .....A.....T........AC....C..T.....T..T..T.........C..C.....T.....T........A..C..C..T.....C..A.....T.
Z.mellis HIS3        .....A.....A........AC.G..C..T..G..T..T..T.........C..C.....T..............A..C..C..T.....C..A.....T.
L.cidri_NCYC_1567    ..T.....T..A..C.....CC.C..T..A.....T..T..T..C........C..C..T..T.....T..C...........T...........A....

                            110       120       130       140       150       160       170       180       190       200
                     ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii HIS3        CATTCAAACAAGCCCTTGGTCAAGTACGTGGTGTGAAAAGATTCGGTTGTGGATTTGCGCCCTTGGATGAAGCGTTATCGCGTGCTGTCGTTGACTTGTC
Z.bisporus HIS3      .......G..G........C.....C...............G..T.....C........A..TC..........T.................G..TC....
Z.lentus HIS3        .T........G..T..A..C.....G.................T....T...G.........TC....C..G..TC....C.....A.....C..T.....
Z.kombuchaensis HIS3 .T........G..T.....C.....G..C...................T...G.....A...C.T.....G..T...........C.....C..T.....
Z.rouxii HIS3        .T........G...T.G..A..G....................T..A.TC..T.........A..A.....G..T.....AA.A.....T..C..TC.A..
Z.mellis HIS3        .T....G.........G..A..G..G..............G..T..C.TC..G.....T..T..A...............AA.A.....A..C..T.....
L.cidri_NCYC_1567    .......GG....T..A...GCT..G........C........T..A..............AC.T..C..G..TC.C........G..T..G..T..A..

                            210       220       230       240       250       260       270       280       290       300
                     ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Z.bailii HIS3        TAATAGACCTTATGCGGTGATTGACCTTGGTTTGAAAAGGGAGAAAATTGGAGACTTATCCTGTGAAATGATACCTCATTTCCTCGAGAGTTTTATTGAA
Z.bisporus HIS3      C..C.....C..C..T..T..A..T..A..C...........A..G.....T.....G..T..C........C..C.....T..A........C..CC.G
Z.lentus HIS3        C..C..G..A..C..A..T......T.A....................C..T...C.G........G...G.C........T..............CC.G
Z.kombuchaensis HIS3 C..C..G.....C..A.....CA..T.G..C........A........C..T...C.G........G.....C........T..............CC.G
Z.rouxii HIS3        C........A...T.T..T.....AT.A..A........A..A.....C..T..T..G..............T..A.....T..A..A.......C....
Z.mellis HIS3        ...C.....A.G.T.T..T.....AT.G..A..A.....A..A.....C..T..T..G.....C........C..A........T..........C....
L.cidri_NCYC_1567    C..C.....A.TC..A...........G..A.....G...........C..T..T.....GACA........T..A.....T..A..A..C..CGC...G

                            310       320
                     ....|....|....|....|.
Z.bailii HIS3        GCTGCAAGGCTGACGGTCCAT
Z.bisporus HIS3      .....C...T....AA.T...
Z.lentus HIS3        .....G..A.....CA.....
Z.kombuchaensis HIS3 .....G..A.....AA.T...
Z.rouxii HIS3        ..A..T..AT....T..T...
Z.mellis HIS3        ..G..C..AT....C......
L.cidri_NCYC_1567    ..A..T..AA.T..TC.....
```

Figure 6.5: ClustalW alignment of the *Zygosaccharomyces species* and *L. cidri* partial *HIS3* gene. Regions for potential *Zygosaccharomyces* genus specific primers have been highlighted.

Chapter 7. References

[1] Barnett, J. A., Payne, R. W., Yarrow, D. and Barnett, L. (2000)
Yeasts: Characteristics and Identification.
Chapter 1 p1-6
Cambridge University Press ISBN: 0521573963.

[2] Rainieri, S., Zambonelli, C. and Kaneko, Y. (2003)
Saccharomyces sensu stricto: Systematics, genetic diversity and evolution
Journal of Bioscience and Bioengineering 96 p1 – 9

[3] Charlier, C., Hart, E., Lefort, A., Ribaud, P., Dromer, F., Denning, D. W. and Lortholary, O. (2006)
Fluconazole for the management of invasive candidiasis: where do we stand after 15 years?
Journal of Antimicrobial Chemotherapy Advance Access, published January 31, 2006.

[4] Barnett, J. A., Payne, R. W., Yarrow, D. and Barnett, L. (2000)
Yeasts: Characteristics and Identification.
Chapter 6 p53-83
Cambridge University Press ISBN: 0521573963.

[5] Kurtzman, C. P. and Fell J. W. (1998)
The Yeasts: A Taxonomic Study, 4th Edition. Chapter 1
Kurtzman, C.P. and Fell, J.W. (editors), Elsevier. ISBN: 0444813128

[6] Moore, R. T. (1998)
The Yeasts: A Taxonomic Study, 4th Edition. Chapter 5
Kurtzman, C.P. and Fell, J.W. (editors), Elsevier. ISBN: 0444813128

[7] Phaff, H. J. (1998).
The Yeasts: A Taxonomic Study, 4th Edition. Chapter 6
Kurtzman, C.P. and Fell, J.W. (editors), Elsevier. ISBN: 0444813128

[8] Gorin, P.A.J. and Spencer, J.F.T. (1970)
Proton magnetic resonance spectroscopy – An aid in identification and chemotaxonomy of yeasts
Advanced Applied Microbiology 13 p25 – 89

[9] Marmur, J. S., Falkow, S. and Mandel, M. (1963)
New approaches to bacterial taxonomy
Annual Review of Microbiology 17 p329 – 372

[10] Storck, R. (1966)
Nucleotide composition of nucleic acids of fungi. II. Deoxyribonucleic acids
Journal of Bacteriology 91 p227 – 230

[11] Stenderup, A. and Bak, A. L. (1968)
Deoxyribonucleic acid base composition of some species within the genus *Candida*
Journal of General Microbiology 52 p231 – 236

[12] Meyer, S.A. and Phaff, H.J. (1969)
Deoxyribonucleic acid base composition in yeasts
Journal of Bacteriology 97 p52 – 56

[13] Meyer, S.A. and Phaff, H.J. (1971)
Taxonomic significance of the DNA base composition in yeasts
In: Ahearn, D.G. Recent Trends In Yeast Research p1 – 29 Georgia State University

[14] Bicknell, J.N. and Douglas, H.C. (1970)
Nucleic acid homologies among species of *Saccharomyces*
Journal of Bacteriology 101 p505 – 512

[15] Price, C.W., Fuson, G.B. and Phaff, H.J. (1978)
Genome comparison in the yeast systematics: delimitation of species within the genera *Schwanniomyces*, *Saccharomyces*, *Debaryomyces* and *Pichia*
Microbiological Reviews 42 p161 – 193

[16] Kurtzman, C. P., Smiley, M. J. and Johnson, C. J. (1980)
Emendation of the genus *Issatchenkia* Kudriavzev and comparison of species by deoxyribonucleic acid reassociation, mating reaction, and ascospore ultrastructure
International Journal of Systematic Bacteriology 30 p503 – 513

[17] Kurtzman, C. P., Smiley, M. J., Johnson, C. J., Wickerham, L. J. and Fuson, G. B. (1980)
Two new and closely related heterothallic species, *Pichia amylophila* and *Pichia mississippiensis*: Characterization by hybridisation and deoxyribonucleic acid reassociation
International Journal of Systematic Bacteriology 30 p208 – 216

[18] Bell, G.I., DeGennare, L.J., Gelfand, D.H., Bishop, R.J., Valenzuela, P. and Rutter, W.J. (1977)
Ribosomal RNA genes of *Saccharomyces cerevisiae*: I. Physical map of the repeating unit and location of the regions coding for 5S, 5.8S, 18S and 25S ribosomal RNAs
The Journal of Biological Chemistry 252 p8118 – 8125

[19] Valenzuela, P., Bell, G. I., Venegas, A., Sewell, E.T., Masiarz, F.R., DeGennare, L.J., Weinberg, F. and Rutter, W.J. (1977)
Ribosomal RNA genes of *Saccharomyces cerevisiae*. II. Physical map and nucleotide sequence of the 5S ribosomal RNA gene and adjacent intergenic regions
The Journal of Biological Chemistry 252 p8126 – 8135

[20] White, T.J., Bruns, T., Lee, S. and Taylor, J. (1990)
Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics
PCR Protocols: Guide to Methods and Applications. Chapter 38, p315 – 322
Innis, M.A., Gelfand, D.H., Sninsky, J.J. and White, T.J. (editors), Academic Press. ISBN: 0123721814

[21] Lachance, M.A. (1990)
Ribosomal DNA spacer variation in the cactophilic yeast *Clavispora opuntiae*
Molecular Biology of Evolution 7 p178 – 193

[22] Magee, B.B., D'Souza, T.M. and Magee, P.T. (1987)
Strain and species identification by restriction fragment length polymorphisms in the ribosomal DNA repeat of *Candida* species
Journal of Bacteriology 169 p1639 – 1643

[23] Clark-Walker, G.D., Hoeben, P., Plazinska, A., Sith, D.K. and Wimmer, E.H. (1987)
Application of mitochondrial DNA analysis to yeast systematics
The Expanding Realm of Yeast-Like Fungi: Proceedings of an international symposium on the perspectives of taxonomy, ecology and phylogeny of yeasts p259 – 266

[24] Peterson, S.W. and Kurtzman, C.P. (1991)
Ribosomal RNA sequence divergence among sibling species of yeasts
Systematic Applied Microbiology 14 p124 – 129

[25] Kurtzman, C.P. (1992)
rRNA sequence comparisons for assessing phylogenetic relationships among yeasts
International Journal of Systematic Bacteriology 42 p1 – 6

[26] Woese, C.R. (2000)
Interpreting the universal phylogenetic tree
Proceedings of the National Academy of Science 97 p8392 – 8396

[27] Fell, J.W., Statzell-Tallman, A., Lutz, M.J. and Kurtzman, C.P. (1992)
Partial rRNA sequences in marine yeasts: a model for identification of marine eukaryotes
Molecular Marine Biology and Biotechnology 1 p175 – 186

[28] James, S.A., Collins, M.D. and Roberts, I.N. (1996)
Use of an rRNA internal transcribed spacer regions to distinguish phylogenetically closely related species of the genera *Zygosaccharomyces* and *Torulaspora*
International Journal of Systematic Bacteriology 40 p189 – 194

[29] Kurtzman, C.P. and Robnett, C.J. (1997)
Identification of clinically important ascomycetous yeasts based on nucleotide divergence in the 5' end of the large-subunit (26S) ribosomal DNA gene
Journal of Clinical Microbiology 35 p1216 – 1223

[30] James, S. A., Cai, J., Roberts, I. N. and Collins, M. D. (1997)
A phylogenetic analysis of the genus *Saccharomyces* based on 18S rRNA gene sequences: Description of *Saccharomyces kunashirensis* sp. nov and *Saccharomyces martiniae* sp. nov.
International Journal of Systematic Bacteriology 47 p453 – 460

[31] Belloch, C., Barrio, E., García, M.D. and Querol, A. (1998)
Phylogenetic reconstruction of the yeast genus *Kluyveromyces*: restriction map analysis of the 5.8S rRNA gene and the two ribosomal internal transcribed spacers
Systematic and Applied Microbiology 21 p266 – 273

[32] Belloch, C., Querol, A., García, M.D. and Barrio, E. (2000)
Phylogeny of the genus *Kluyveromyces* inferred from the mitochondrial cytochrome-c oxidase II gene
International Journal of Systematic and Evolutionary Microbiology 50 p405 – 416

[33] Cappa F. and Cocconcelli P.S. (2001)
Identification of fungi from dairy products by means of 18S rRNA analysis
International Journal of Food Microbiology 69 p157-160

[34] Kurtzman, C.P. and Robnett, C.J. (1998)
Identification and phylogeny pf ascomycetous yeasts from analysis of nuclear large subunit (26S) ribosomal DNA partial sequences
Antonie van Leeuwenhoek 73 p331 – 371

[35] Rokas, A., Williams, B.L., King, N. and Carroll, S.B. (2003)
Genome-scale approaches to resolving incongruence in molecular phylogenies
Nature 425 p798 – 804

[36] Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1998)
Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase
Science 239 p487-91

[37] Richard G-F., Hennequin C., Thierry A. and Dujon B. (1999)
Trinucleotide repeats and other microsatellites in yeasts
Research in Microbiology 150 p589-602

[38] Andrighetto, C., Psomas, E., Tzanetakis, N., Suzzi, G. and Lombardi, A. (2000)
Randomly amplified polymorphic DNA (RAPD) PCR for the identification of yeasts isolated from dairy products
Letters in Applied Microbiology 30 p5 – 9

[39] Baleiras Couto, M.M., Hartog, B.J., Huis in't Veld, J.H.J., Hofstra, H. and van der Vossen, J.M.B.M. (1996)
Identification of spoilage yeasts in a food production chain by microsatellite polymerase chain reaction fingerprinting
Food Microbiology 13 p59 – 67

[40] Malgoire, J. Y., Bertout, S., Renaud, F., Bastide, J. M. and Mallié, M. (2005)
Typing of *Saccharomyces cerevisiae* clinical strain by using microsatellite sequence polymorphism
Journal of Clinical Microbiology 43 p1133 – 1137

[41] Pafetti, D., Barerio, C., Casalone, E., Cavalieri, D., Fani, R., Fia, G., Mori, E. and Polsinelli, M. (1995)
DNA fingerprinting by random amplified polymorphic DNA and restriction fragment length polymorphism is useful for yeast typing
Research in Microbiology 146 p587 – 594

[42] Baleiras Couto, M.M., Eijsma, B., Hofstra, H., Huis in't Veld, J.H.J. and van der Vossen, J.M.B.M. (1996)
Evaluation of molecular typing techniques to assign genetic diversity among *Saccharomyces cerevisiae* strains
Applied and Environmental Microbiology 62 p41-46

[43] González Techera A., Jubany S., Carrau F.M. and Gaggero C. (2001)
Differentiation of industrial wine yeast strains using microsatellite markers
Letters in Applied Microbiology 33 p71-75

[44] Pérez M.A., Gallego F.J., Martínez I. and Hidalgo P. (2001)
Detection, distribution and selection of microsatellites (SSRs) in the genome of the yeast *Saccharomyces cerevisiae* as molecular markers
Letters in Applied Molecular Biology 33 p461-466

[45] Caruso, M., Capece, A., Salzano, G. and Romano, P. (2002)
Typing of *Saccharomyces cerevisiae* and *Kloeckera apiculata* strains from Aglianico wine
Letters in Applied Microbiology 34 p323 – 328

[46] Pearson, B.M., Carter, A.,T., Furze, J.M. and Roberts, I.N. (1995)
A novel approach for discovering retrotransposons: Characterization of a long terminal repeat element in the spoilage yeast *Pichia membranefaciens* and its use in strain identification
International Journal of Systematic Bacteriology 45 p386 – 389

[47] Sampaio, P., Gusmão, L., Correia, A., Alves, C., Rodrigues, A. G., Pina-Vaz, C., Amorim, A. and Pais, C. (2005)
New microsatellite multiplex PCR for Candida albicans strain typing reveals microevolutionary changes
Journal of Clinical Microbiology 43 p3869 – 3876

[48] Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kulper, M and Zabeau, M. (1995)
AFLP: A new technique for DNA fingerprinting
Nucleic Acids Research 23 p4407 – 4414

[49] de Barros Lopes, M., Rainieri, S., Henschke, P. A. and Langridge, P. (1999)
AFLP fingerprinting for analysis of yeast genetic variation
International Journal of Systematic Bacteriology 49 p915 – 924

[50] Flores Berrios, E. P., Alba González, J. F., Arrizon Gaviño, J. P., Romano, P., Capece, A. and Gschaedler Mathis, A. (2005)
The uses of AFLP for detecting DNA polymorphism, genotype identification and genetic diversity between yeasts isolated from Mexican agave-distilled beverages and from grape musts
Letters in Applied Microbiology 41 p147 - 152

[51] Beltran, G., Torija, M.J., Novo, M., Ferrer, N., Poblet, M., Guillamón, J.M., Rosès, N. and Mas, A. (2002)
Analysis of yeast populations during alcoholic fermentation: A six year follow-up study
Systematic Applied Microbiology 25 p287 – 293

[52] Fernández, M.T., Ubeda, J.F. and Briones, A.I. (1999)
Comparative study of non-*Saccharomyces* microflora of musts in fermentation, by physiological and molecular methods
FEMS Microbiology Letters 173 p223 – 229

[53] Fernández-Espinar, M.T, López, V., Ramón, D., Bartra, E. and Querol, A. (2001)
Study of the authenticity of commercial wine yeast strains by molecular techniques
International Journal of Food Microbiology 70 p1 – 10

[54] Granchi, L., Bosco, M., Messini, A. and Vincenzini, M. (1999)
Rapid detection and quantification of yeast species during spontaneous wine fermentation by PCR-RFLP analysis of the rDNA ITS region
Journal of Applied Microbiology 87 p949-956

[55] Guillamón, J.M., Sabaté, J., Barrio, E., Cano, J. and Querol, A. (1998)
Rapid identification of wine yeast species based on RFLP analysis of the ribosomal internal transcribed spacer (ITS) region
Archaeological Microbiology 196 p387 – 392

[56] Pramateftaki P.V., Lanaridis P. and Typas M.A. (2000)
Molecular identification of wine yeasts at species or strain level: a case study with strains from two vine-growing areas of Greece
Journal of Applied Microbiology 89 p236-248

[57] Schuller, D., Alves, H., Dequin, S. and Casal, M. (2005)
Ecological survey of *Saccharomyces cerevisiae* strains from vineyards in the Vinho Verde region of Portugal
FEMS Microbiology Ecology 51 p167 – 177

[58] Caggia C., Restuccia C., Pulvirenti A. and Giudici P. (2001)
Identification of *Pichia anomala* isolated from yoghurt by RFLP of the ITS region
International Journal of Food Microbiology 71 p71-73

[59] Guillamón J.M., Sánchez I. and Huerta T. (1997)
Rapid characterization of wild and collection strains of the genus *Zygosaccharomyces* according to mitochondrial DNA patterns
FEMS Microbiology Letters 147 p267-272

[60] Gupta, A.K., Kohli, Y. and Summerbell, R.C. (2000)
Molecular differentiation of seven *Malassezia* species
Journal of Clinical Microbiology 38 p1869 – 1875

[61] Peterson, K.M. Møller, P.L. and Jespersen, L. (2001)
DNA typing methods for differentiation of *Debaryomyces hansenii* strains and other yeasts related to surface ripened cheeses
International Journal of Food Microbiology 69 p11 – 24

[62] Smole Možina, S., Dlauchy, D. and Raspor, P. (1997)
Identification of *Saccharomyces sensu stricto* and *Torulaspora* yeasts by PCR ribotyping
Letters in Applied Microbiology 24 p311 – 315

[63] Yamagishi, H., Otsuta, Y., Funahashi, W., Ogata, T. and Sakai, K. (1999)
Differentiation between brewing and non-brewing yeasts using a combination of PCR and RFLP
Journal of Applied Microbiology 86 p505 – 513

[64] Tornai-Lehoczki J. and Dlauchy D. (2000)
Delimitation of brewing yeast strains using different molecular techniques
International Journal of Food Microbiology 62 p37-45

[65] Muyzer G. (1999)
DGGE/TGGE a method for identifying genes from natural ecosystems
Current Opinions in Microbiology 2 p317-322

[66] Cocolin L., Bisson L.F. and Mills D.A. (2000)
Direct profiling of the yeast dynamics in wine fermentations
FEMS Microbiology Letters 189 p81-87

[67] Nielsen, D. S., Honholt, S., Tano-Debrah, K. and Jespersen, L (2005)
Yeast populations associated with Ghanaian cocoa fermentations analysed using denaturing gradient gel electrophoresis (DGGE)
Yeast 22 p271 – 84

[68] Meroth, C., Hammes, W. P. and Hertel, C. (2003)
Identification and population dynamics of yeasts in sourdough fermentation processes by PCR-denaturing gradient gel electrophoresis
Applied and Environmental Microbiology 69 p7453 – 7461

[69] Manzano, M., Cocolin, L., Iacumin, L., Cantoni, C. and Comi, G. (2005)
A PCR-TGGE (temperature gradient gel electrophoresis) technique to assess differentiation among enological *Saccharomyces cerevisiae* strains
International Journal of Food Microbiology 101 p333 – 339

[70] Tominaga, T. (2004)
Rapid identification of pickle yeasts by fluorescent PCR and microtemperature-gradient gel electrophoresis
FEMS Microbiology Letters 238 p43 – 48

[71] Versavaud, A. and Hallet, J.N. (1995)
Pulsed-field gel electrophoresis combined with rare-cutting endonucleases for strain differentiation of *Candida famata*, *Kloeckera apiculata* and *Schizosaccharomyces pombe* with chromosome number and size estimation of the two former
Systematic Applied Microbiology 18 p303 – 309

[72] Gomes L.H., Duarte K.M.R., Argueso J.L., Echeverrigaray S. and Tavares F.C.A. (2000)
Methods for yeast characterization from industrial products
Food Microbiology 17 p217-223

[73] Schütz M. and Gafner J. (1994)
Dynamics of the yeast strain population during spontaneous alcoholic fermentation determined by CHEF gel electrophoresis
Letters in Applied Microbiology 19 p253-257

[74] Giudici P., Caggia C., Pulvirenti A., Zambonalli C. and Rainieri S. (1998)
Electrophoretic profile of hybrids between cryotolerant and non-cryotolerant *Saccharomyces* strains
Letters in Applied Microbiology 27 p31-34

[75] Yamazaki, M., Kurtzman, C. P. and Sugiyama, J. (1998)
The Yeasts: A Taxonomic Study, 4th Edition. Chapter 7
Kurtzman, C.P. and Fell, J.W. (editors), Elsevier. ISBN: 0444813128

[76] Perry-O'Keefe H., Stender H., Broomer A., Oliveira K., Coull J. and Hyldig-Nielson J.J. (2001)
Filter-based PNA *in situ* hybridization for rapid detection, identification and enumeration of specific micro-organisms
Journal of Applied Microbiology 90 p180-189

[77] Sancho T., Giménez-Jurado G., Malfieto-Ferriera M. and Loureiro V. (2000)
Zymological indicators: a new concept applied to the detection of potential spoilage yeast species associated with fruit pulps and concentrates
Food Microbiology 17 p613-624

[78] Arias, C.R., Burns, J.K., Friedrich, L.M., Goodrich, R.M. and Parish, M.E. (2002)
Yeast species associated with orange juice: Evaluation of different identification methods
Applied and Environmental Microbiology 68 p1955 – 1961

[79] Antunovics, Z., Irinyi, L. and Sipiczki, M. (2005)
Combined application of methods to taxonomic identification of *Saccharomyces* strains in fermenting botryized grape must
Journal of Applied Microbiology 98 p971 – 979

[80] de Barros Lopes M., Soden A., Henschke P.A. and Langridge P. (1996)
PCR differentiation of commercial yeast strains using intron splice site primers
Applied and Environmental Microbiology 62 p4514 – 4520

[81] de Barros Lopes M., Soden A., Martens A.L., Henschke P.A. and Langridge P. (1998)
Differentiation and species identification of yeasts using PCR
International Journal of Systematic Bacteriology 48 p279-286

[82] Egli, C.M. and Henick-Kling, T. (2001)
Identification of *Brettanomyces / Dekkera* species based on polymorphism in the rRNA internal transcribed spacer region
American Journal of Enology and Viticulture 52 p241 – 247

[83] Sugita, T., Nishikawa, A. and Shinoda, T. (1998)
Identification of *Trichosporon asahii* by PCR based on sequences of the internal transcribed spacer regions
Journal of Clinical Microbiology 36 p2742 – 2744

[84] Casey, G.D. and Dobson, A.D.W. (2003)
Molecular detection of *Candida krusei* contamination in fruit juice using the citrate synthase gene *cs1* and a potential role for this gene in the adaptive response to acetic acid
Journal of Applied Microbiology 95 p13 – 22

[85] Donnelly, S.M., Sullivan, D.J., Shanley, D.B. and Coleman, D.C. (1999)
Phylogenetic analysis and rapid identification of *Candida dubliniensis* based on analysis of *ACT1* intron and exon sequences
Microbiology 145 p1871 – 1882

[86] Zapparoli G., Torriani S., Pesente P. and Dellaglio F. (1998)
Design and evaluation of malolactic enzyme gene targeted primers for rapid identification of *Oenococcus oeni* in wine
Letters in Applied Microbiology 27 p243-246

[87] Fell, J.W. (1993)
Rapid identification of yeast species using three primers in a polymerase chain reaction
Molecular Marine Biology and Biotechnology 2 p174 – 180

[88] Mannarelli, B.M. and Kurtzman, C.P. (1998)
Rapid identification of *Candida albicans* and other human pathogenic yeasts by using short oligonucleotides in a PCR
Journal of Clinical Microbiology 36 p1634 – 1641

[89] Kosse, D., Seiler, H., Amann, R., Ludwig, W. and Scherer, S. (1997)
Identification of yoghurt spoiling yeasts with 18S rRNA-targeted oligonucleotide probes
Systematic Applied Microbiology 20 p568 – 480

[90] Jaeger, E. E. M., Carroll, N. M., Choudhury, S., Dunlop, A. A. S., Towler, H. M. A., Matheson, M. M., Adamson, P., Okhravi, N. and Lightman, S. (2000)
Rapid detection and identification of *Candida*, *Aspergillus*, and *Fusarium* species in ocular samples using nested PCR
Journal of Clinical Microbiology 38 p2902 – 2908

[91] Bougnoux, M.-E., Dupont, C., Mateo, J., Saulnier, P., Faivre, V., Payen, D. and Nicolas-Chanoine, M.-H. (1999)
Serum is more suitable than whole blood for diagnosis of systemic Candidiasis by nested PCR
Journal of Clinical Microbiology 37 p925 – 930

[92] Bleve, G., Rizzotti, L., Dellaglio, F. and Torriani, S. (2003)
Development of reverse transcription (RT)-PCR and real-time RT-PCR assays for rapid detection and quantification of viable yeasts and molds contaminating yogurts and pasteurised food products
Applied and Environmental Microbiology 69 p4116 – 4122

[93] Brinkman, N.E., Haugland, R.A., Wymer, L.J., Byappanahalli, M., Whitman, R.L. and Vesper, S.J. (2003)
Evaluation of a rapid, quantitative real-time PCR methods for enumeration of pathogenic *Candida* cells in water
Applied and Environmental Microbiology 69 p1775 – 1782

[94] Casey, G.D. and Dobson, A.D.W. (2004)
Potential of using real-time PCR-based detection of spoilage yeast in fruit juices – a preliminary study
International Journal of Food Microbiology 91 p327 – 335

[95] Jothikumar, N. and Griffiths, M. W. (2002)
Rapid detection of *Escherichia coli* O157:H7 with multiplex real-time PCR assays
Applied and Environmental Microbiology 68 p3169 – 3171

[96] Misonne, M-C. and Hoet, P. P. (1998)
Species-specific plasmid sequences for PCR identification of the three species of *Borrelia burgdorferi sensu lato* involved in Lyme disease
Journal of Clinical Microbiology 36 p269 – 272

[97] Hein, I., Lehner, A., Rieck, P., Klein, K., Brandl, E. and Wagner, M. (2001)
Comparison of different approaches to quantify *Staphylococcus aureus* cells by real-time quantitative PCR and application of this technique for examination of cheese
Applied and Environmental Microbiology 67 p3122 – 3126

[98] Logan, J. M. J., Edwards, K. J., Saunders, N. A. and Stanley, J. (2001)
Rapid identification of *Campylobacter* spp. by melting peak analysis of biprobes in real-time PCR
Journal of Clinical Microbiology 39 p2227 – 2232

[99] Mommert, S., Gutzmer, R., Kapp, A. and Werfel, T. (2001)
Sensitive detection of *Borrelia burgdorferi sensu lato* DNA and differentiation of *Borrelia* species by LightCycler PCR
Journal of Clinical Microbiology 39 p2663 – 3667

[100] Nogva, H. K., Rudi, K., Naterstad, K., Holck, A and Lillehaug, D. (2000)
Application of 5'-nuclease PCR for quantitative detection of *Listeria monocytogenes* in pure cultures, water, skim milk and unpasteurized whole milk
Applied and Environmental Microbiology 66 p4266 – 4271

[101] Abliz, P., Fukushima, K., Takizawa, K., Nieda, N., Miyaji, M. and Nishimura, K. (2003)
Rapid identification of the genus *Fonsecaea* by PCR with specific oligonucleotide primers
Journal of Clinical Microbiology 41 p873 – 876

[102] Bell, K. S., Roberts, J., Verrall, S., Cullen, D. W., Williams, N. A., Harrison, J. G., Toth, I. K., Cooke, D. E. L., Duncan, J. M. and Claxton, J. R. (1999)
Detection and quantification of *Spongospora subterranea* f. sp. *subterranea* in soils and on tubers using specific PCR primers
European Journal of Plant Pathology 105 p905 – 915

[103] Salazar, O., Julian, M. C. and Rubio, V. (2000)
Primers based on specific rDNA-ITS sequences for PCR detection of *Rhizoctonia solani*, *R. solani* AG 2 subgroups and ecological types, and binucleate *Rhizoctonia*
Mycology Research 104 p281 – 285

[104] Suga, H., Hasegawa, T., Mitsui, H., Kageyama, K. and Hyakumachi, M. (2000)
Phylogenetic analysis of the phytopathogenic fungus *Fusarium solani* based on the rDNA-ITS region
Mycology Research 104 p1175 – 1183

[105] Wang, P. H., Wang, Y. T. and White, J. G. (2001)
Species-specific PCR primers for *Pythium* developed from ribosomal ITS1 region
Letters in Applied Microbiology 37 p127 – 132

[106] Gutzmer, R., Mommert, S., Küttler, U., Werfel, T. and Kapp, A. (2004)
Rapid identification and differentiation of fungal DNA in dermatological specimens by LightCycler PCR
Journal of Medical Microbiology 53 p1207 – 1214

[107] Stratford, M., Hofman, P.D. and Cole, M.B. (2000)
Fruit juices, fruit drinks and soft drinks
The Microbiological Safety and Quality of Food. Volume I. Chapter 31. 836 – 869
Lund, B.M., Baird-Parker, T.C. and Gould, G.W. (editors), Aspen Publishers Inc. ISBN: 0834213230

[108] NCBI homepage: http://www.ncbi.nlm.nih.gov/

[109] Hall, T.A. (1999)
Bioedit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT
Nucleic Acids Symposium 41 p95 – 98

[110] Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994)
CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice
Nucleic Acids Research 22 p4673 – 4680

[111] Kumar, S., Tamura, K. and M Nei (2004)
MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment.
Briefings in Bioinformatics 5 p150 – 163

[112] Smith, M.Th. (1998)
Brettanomyces Kufferath & van Laer
The Yeasts: A Taxonomic Study, 4th Edition. Chapter 63. 450 – 453
Kurtzman, C.P. and Fell, J.W. (editors), Elsevier. ISBN: 0444813128

[113] Smith, M.Th. (1998)
Dekkera van der Walt
The Yeasts: A Taxonomic Study, 4^{th} Edition. Chapter 27. 174 – 177
Kurtzman, C.P. and Fell, J.W. (editors), Elsevier. ISBN: 0444813128

[114] Fugelsang, K.C. (1997)
Wine Microbiology
The Chapman and Hall Enology Library. International Thompson Publishing.

[115] Van der Walt, J.P. (1964)
Dekkera, a new genus of the *Saccharomycetaceae*
Antonie van Leeuwenhoek 30 p273 – 280

[116] Lee, F.W. and Jong, S.C. (1985)
Dekkera claussenii sp. nov., The perfect state of *Brettanomyces claussenii*
Mycotaxon 23 p275 – 278

[117] Lee, F.W. and Jong, S.C. (1986)
The new species *Dekkera abstinens*, teleomorph of *Brettanomyces abstinens*
Mycologia 78 p150 – 151

[118] Lee, F.W. and Jong, S.C. (1986)
New species *Dekkera custersiana* and *D. lambica*, teleomorphs of *Brettanomyces*
Mycotaxon 25 p455 – 460

[119] Lee, F.W. and Jong, S.C. (1986)
The new species *Dekkera naardenesis*, teleomorph of *Brettanomyces naardenensis*
Mycotaxon 25 p147 – 152

[120] Hoeben, P. and Clark-Walker, G.D. (1986)
An approach to yeast classification by mapping mitochondrial DNA from *Dekkera / Brettanomyces* and *Eeniella* genera
Current Genetics 10 p371 – 379

[121] Smith, M.Th., Yamazaki, M. and Poot, G.A. (1990)
Dekkera, *Brettanomyces* and *Eeniella*: Electrophoretic comparison of enzymes and DNA – DNA homology
Yeast 6 p299 – 310

[122] Molina, F.I., Shen, P. and Jong, S.C. (1993)
Validation of the species concept in the genus *Dekkera* by restriction analysis of genes coding for rRNA
International Journal of Systematic Bacteriology 43 p32 – 35

[123] Boekhout, T., Kurtzman, C.P., O'Donnell, K. and Smith, M.Th. (1994)
Phylogeny of the yeast genera *Hanseniaspora* (Anamorph *Kloeckera*), *Dekkera* (anamorph *Brettanomyces*), and *Eeniella* as inferred from partial 26S ribosomal DNA nucleotide sequences
International Journal of Systematic Bacteriology 44 p781 – 786

[124] Yamada, Y., Matsuda, M., Maeda, K. and Mikata, K. (1994)
The phylogenetic relationships of species of the genus *Dekkera* van der Walt based on the partial sequences of 18S and 26S ribosomal RNAs (Saccharomycetaceae)
Bioscience, Biotechnology and Biochemistry 58 p1803 – 1808

[125] Entrez nucleotide sequence search
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Nucleotide

[126] NCBI BLAST
http://www.ncbi.nlm.nih.gov/BLAST/

[127] Kurtzman, C.P. (1998)
Zygosaccharomyces Barker
The Yeasts: A Taxonomic Study, 4^{th} Edition. Chapter 57. p424 – 432
Kurtzman, C.P. and Fell, J.W. (editors), Elsevier. ISBN: 0444813128

[128] Steels, H., Bond, C.J., Collins, M.D., Roberts, I.N., Stratford, M. and James, S.A. (1999)
Zygosaccharomyces lentus sp. nov., a new member of the yeast genus *Zygosaccharomyces* Barker
International Journal of Systematic Bacteriology 49 p319 – 327

[129] Steels, H., James, S.A., Roberts, I.N. and Stratford, M. (1999)
Zygosaccharomyces lentus: a significant new osmophilic, preservative-resistant spoilage yeast, capable of growth at low temperatures
Journal of Applied Microbiology 87 p520 – 527

[130] Kurtzman, C.P., Robnett, C.J. and Basehoar-Powers, E. (2001)
Zygosaccharomyces kombuchaensis: a new ascosporogenous yeast from 'Kombucha tea'
FEMS Yeast Research 1 p133 – 138

[131] Steels, H., James, S.A., Bond, C.J., Roberts, I.N. and Stratford, M. (2002)
Zygosaccharomyces kombuchaensis: the physiology of a new species related to the spoilage yeasts *Zygosaccharomyces lentus* and *Zygosaccharomyces bailii*
FEMS Yeast Research 2 p113 – 121

[132] Kurtzman, C.P. (1998)
Torulaspora Lindner
The Yeasts: A Taxonomic Study, 4^{th} Edition. Chapter 51. p404 – 408
Kurtzman, C.P. and Fell, J.W. (editors), Elsevier. ISBN: 0444813128

[133] Kurtzman, C. P. (2003)
Phylogenetic circumscription of *Saccharomyces*, *Kluyveromyces* and other members of the Saccharomycetaceae, and the proposal of the new genera *Lachancea*, *Nakaseomyces*, *Naumovia*, *Vanderwaltozyma* and *Zygotorulaspora*
FEMS Yeast Research 4 p233 – 245

[134] James, S. A., Bond, C. J. and Roberts, I. N. (2005)
Molecular evidence for the existence of natural hybrids in the genus *Zygosaccharomyces*
FEMS Yeast Research 5 p747 – 755

[135] Fundyga, R. E., Kuykendall, R.J., Wang, L.-Y. and Dubourdieu, D. (2004)
Evidence for aneuploidy and recombination in the human commensal east *Candida parapsilosis*
Infection, Genetics and Evolution 4 p37 – 907

[136] Boekhout, T., Theelan, B., Diaz, M., Fell, J. W., Hopp, W. C. J., Abeln, E. C. A., Dromer, F. and Meyer, W. (2001)
Hybrid genotypes in the pathogenic yeast *Cryptococcus neoformans*
Microbiology 147 p891 – 907

[137] Vaughan-Martini, A. and Martini, A. (1998)
Saccharomyces Meyen ex Reess
The Yeasts: A Taxonomic Study, 4th Edition. Chapter 44. p358 – 371
Kurtzman, C.P. and Fell, J.W. (editors), Elsevier. ISBN: 0444813128

[138] Wolfe, K. H. and Shields, D. C. (1997)
Molecular evidence for an ancient duplication of the entire yeast genome
Nature 387 p708 – 713

[139] Wolfe, K. (2004)
Evolutionary genomics: Yeasts accelerate beyond BLAST
Current Biology 14 pR392 – R394

[140] Ochman, H., Daubin, V. and Lerat, E. (2005)
A bunch of fun-guys: the whole-genome view of yeast evolution
Trends in Genetics 21 p1 – 3

[141] Langkjær, R. B., Cliften, P. F., Johnston, M and Piškur, J. (2003)
Yeast genome duplication was followed by asynchronous differentiation of duplicated genes
Nature 421 p848 – 852

[142] Conant, G. C. and Wagner, A. (2003)
Asymmetric sequence divergence of duplicate genes
Genome Research 13 p2052 – 2058

[143] Masneuf, I., Hansen, J., Groth, C., Piskur, J. and Dubourdieu, D. (1998)
New hybrids between *Saccharomyces sensu stricto* yeast species found among wine and cider production strains
Applied and Environmental Microbiology 64 p3887 – 3892

[144] Sebastiani, F., Barberio, C., Casalone, E., Cavalieri, D. and Polsinelli, M. (2002)
Crosses between *Saccharomyces cerevisiae* and *Saccharomyces bayanus* generate fertile hybrids
Research in Microbiology 153 p53 – 58

[145] Naumov, G.I., James, S.A., Naumova, E.S., Louis, E.J. and Roberts, I.N. (2000)
Three new species in the *Saccharomyces sensu stricto* complex: *Saccharomyces cariocanus*, *Saccharomyces kudriavzevii* and *Saccharomyces mikatae*
International Journal of Systematic and Evolutionary Microbiology 50 p1931 – 1942

[146] Vaughan-Martini, A. and Kurtzman, C. P. (1985)
Deoxyribonucleic acid relatedness among species of *Saccharomyces sensu stricto*
International Journal of Systematic Bacteriology 35 p508 – 511

[147] Nguyen, H.-V. and Gaillardin, C. (1997)
Two subgroups within the *Saccharomyces bayanus* species evidenced by PCR amplification and restriction polymorphism of the non-transcribed spacer 2 in the ribosomal DNA unit
Systematic and Applied Microbiology 20 p286 – 294

[148] Rainieri, S., Zambonelli, C., Hallsworth, J. E., Pulvirenti, A. and Giudici, P. (1999)
Saccharomyces uvarum, a distinct group within *Saccharomyces sensu stricto*
FEMS Microbiology Letters 1777 p1777 – 185

[149] Pulvirenti, A., Nguyen, H.-V., Caggia, C., Giudici, P., Rainieri, S. and Zambonelli, C. (2000)
Saccharomyces uvarum, a proper species with the *Saccharomyces sensu stricto*
FEMS Microbiology Letters 192 p 191 – 196

[150] Nguyen, H.-V., Lépingle, A. and Gaillardin, C. (2000)
Molecular typing demonstrates homogeneity of *Saccharomyces uvarum* and reveals the existence of hybrids between *S. uvarum* and *S. cerevisiae*, including the *S. bayanus* type strain CBS 380
Applied Microbiology 23 p71 – 85

[151] Lodder, J. and Kreger-van Rij, N. J. W. (1952)
The Yeasts, a taxonomic study
Amsterdam: North Holland

[152] Casaregola, S., Nguyen, H.-V., Lapathitis, G., Kotyk, A. and Gaillardin, C. (2001)
Analysis of the constitution of the beer yeast genome by PCR, sequencing and subtelomeric sequence hybridization
International Journal of Systematic and Evolutionary Biology 51 p1607 – 1618

[153] Nguyen, H.-V. and Gaillardin, C. (2005)
Evolutionary relationships between the former species *Saccharomyces uvarum* and the hybrids *Saccharomyces bayanus* and *Saccharomyces pastorianus*; reinstatement of *Saccharomyces uvarum* (Beijerinck) as a distinct species
FEMS Yeast Research 5 p471 – 483

[154] *Saccharomyces* genome database website: http://www.yeastgenome.org/

[155] Kellis, M., Patterson, N., Endrizzi, M., and Lander, E. S. (2003)
Sequencing and comparison of yeast species to identify genes and regulatory elements
Nature 423 p241 – 254

[156] Cliften, P., Sudarsanam, P., Desikan, A., Fulton, L., Fulton, B., Majors, J., Waterston, R., Cohen, B. A. and Johnston, M. (2003)
Finding functional features in *Saccharomyces* genomes by phylogenetic footprinting
Science 301 p71 – 76

[157] *S. kudriavzevii* sequence data source: (Washington University in St Louis, Department of Genetics)
http://www.genetics.wustl.edu/saccharomycesgenomes/download.html

[158] Sniegowski, P. D., Dombrowski, P. G. and Fingerman, E. (2002)
Saccharomyces cerevisiae and *Saccharomyces paradoxus* coexist in a natural woodland site in North America and display different levels or reproductive isolation from European conspecifics
FEMS Yeast Research 1 p299 – 306

[159] Lachance, M. A., Danial, H. M., Meyer, S., Prasad, G. S., Gautam, S. P. and Boundy-Mills, K. (2003)
The D1 / D2 domain of the large-subunit rDNA of the yeast species Clavispora lusitaniae is unusually polymorphic
FEMS Yeast Research 4 p253 - 258

[160] Kwok, S., Kellogg, D. E., McKinney, N., Spasic, D., Levenson, C. and Sninsky, J. J. (1990)
Effects of primer – template mismatches in the polymerase chain reaction: Human immunodeficiency virus type 1 model studies
Nucleic Acids Research 18 p999 – 1005

[161] Huang, M.-M., Arnheim, N. and Goodman, M. F. (1992)
Extension of base mispairs by TAP DNA polymerase: implications for single nucleotide discrimination in PCR
Nucleic Acids Research 20 p4567 – 4573

[162] Matsuhashi, M. (1996)
Technical report: Part 1. Basic requirements for designing optimal oligonucleotide probe sequences
Journal of Clinical Laboratory Analysis 10 p277 – 284

[163] Allawi, T. H. and SantaLucia, J. (1997)
Thermodynamics and NMR of internal G.T mismatches in DNA
Biochemistry 36 p10581 – 10594

[164] Allawi, T. H. and SantaLucia, J. (1998)
Nearest neighbor thermodynamic parameters for internal G.A mismatches in DNA
Biochemistry 37 p2170 – 2179

[165] Allawi, T. H and SantaLucia, J. (1998)
Nearest-neighbor thermodynamics of internal A.C mismatches in DNA: Sequence dependence and pH effects
Biochemistry 37 p9435 – 9444

[166] Peyret, N., Seneviratne, A., Allawi, H. T. and SantaLucia, J. (1999)
Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A.A, C.C, G.G and T.T mismatches
Biochemistry 38 p3468 – 3477

[167] Integrated DNA technologies, Oligoanalyzer 3.0
http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/Default.aspx

[168] Pryce, T. M., Palladino, S., Price, D. M., Gardam, D. J., Campbell, P. B., Christiansen, K. J. and Murray, R. J. (2006)
Rapid identification of fungal pathogens in BacT/ALERT, BACTEC and BBL MGIT media using polymerase chain reaction and DNA sequencing of the internal transcribed spacer regions
Diagnostic Microbiology and Infectious Disease IN PRESS